CONFERENCE PROCEEDINGS

HUFLIT INTESOL INTERNATIONAL CONFERENCE 2026

Towards English as a Second Language in Vietnam: Pedagogical, Policy, and Practice Perspectives



07 August 2026 HUFLIT

STEERING COMMITTEE

1) Dr. Nguyen Anh Tuan - Vice-Chairman of Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT).
2) Asc. Prof. Dr. Tran Manh Ha - President of HUFLIT.
3) Asc. Prof. Dr. Nguyen Ngoc Vu - Vice President of HUFLIT.
4) Dr. Nguyen Thuy Hong Van - Managing Director of Veducon, Director of INTESOL Vietnam.

ORGANIZING COMMITTEE

No.	Fullname	Title - Organaztion	Position in the Organizing Committee
1	Dr. Nguyen Anh Tuan	Vice-Chairman of HUFLIT Council	Chairman
2	Asc. Prof. Dr. Tran Manh Ha	President, HUFLIT	Deputy Chairman
3	Asc. Prof. Dr. Nguyen Ngoc Vu	Vice President, Director of TARI, HUFLIT	Permanent Deputy Chairman
4	Dr. Nguyen Thuy Hong Van	Managing Director of Veducon, Director of INTESOL Vietnam	Deputy Chairman
5	Asc. Prof. Dr. Nguyen Thanh Binh	Deputy Head of Department of Postgraduate Education - Science and Technology, HUFLIT	Permanent Member
6	Ms. Lương Thị Hồng Cẩm	Head of Communications - Public Relations, Head of Marketing Department at Veducon & Intesol Vietnam	Member
7	Dr. Thai Thi Cam Trang	Vice Director of EDC, Hanoi National University of Education	Member
8	Dr. Do Thi Phuong Mai	Head of the English Language Department - Faculty of English, Academic Director of EDC, HNUE	Member
9	Asc. Prof. Dr. Tran Van Lang	Head of Department of Science Journal, HUFLIT	Member
10	Bui Thi Thanh Truc, M.A.	Dean of Faculty of Foreign Languages, HUFLIT	Member
11	Pham Phuc Huy, MBA.	Head of Department of Planning and Finance, HUFLIT	Member
12	Nguyen Thi Thieu, MAFM.	Head of Department of External Relations, HUFLIT	Member
13	Nguyen Duy Trung, MBA.	Deputy Director of Center of Communications - Event Management, HUFLIT	Member

14	Nguyen Quoc Hung, MHrm.	Head of Department of Human Resources and Administrative Affairs, HUFLIT	Member
15	Huynh Thi Nhi, M.A.	Head of Department of Facilities Purchasing, HUFLIT	Member
16	Mr. Nguyen Minh Son	Director of Center for Facilities Operations, HUFLIT	Member
17	Dinh Hong Van, MBA.	Head of Department of Student Affairs, HUFLIT	Member
18	Dang Quoc Tuan, MBA	Director of Center of Information Technology & Data, HUFLIT	Member

PROCEEDINGS EDITORIAL BOARD

Chair:

1) Asc. Prof. Dr. Nguyen Ngoc Vu -Vice President of HUFLIT

Committee Members:

1) Le Thi Thao, MBA. - HUFLIT
2) Mr. Tchen Van An - HUFLIT
3) Nguyen Thi My Hang, MBA. - HUFLIT

TABLE OF CONTENT

INTRODUCTION xviii

#1 The Impacts of Various Teaching Methods and Assessments Applied in Business Communication Course 19

Thái Mỹ Linh 19

#2 EVALUATING PROGRAM LEARNING OUTCOME CHIEVEMENT THROUGH A DATA-INFORMED OUTCOME-BASED EDUCATION APPROACH: EVIDENCE FROM ENGLISH-MAJORED UNDERGRADUATES AT HUFLIT UNIVERSITY 36

Tran Trung Hieu , Tran Kim Tuyen 36

#3 Transinterpretation Pedagogy and Application of Artificial Intelligence for better motivation: Empirical Evidence from TMT Exercises 54

Nguyen Duc Chau 54

#7 Translation-majored Students' ChatGPT Usage Pattern and Translation Self-Efficacy 74

Quan Ngo, Chau Nguyen 74

#15 From Classroom to Real Life: First-Year Students' English-Speaking Use across Learning Contexts 94

Nguyen Thi Ngoc Chinh 94

#18 Exploring University Students' Use of AI Tools in Academic Writing at HUFLIT 109

Phan Ngoc Phuong Vy 109

#21 EFL-TO-ESL REPOSITIONING THROUGH TEACHER DEVELOPMENT: IMPLICATIONS FOR INTERNATIONALIZED EDUCATION AND WORKPLACES 124

Nguyễn Hưng Bình 124

#23 Semiotics and Translanguaging Approaches to Inclusion and Critical Thinking in Generative AI Adoption 134

Fati Osuman 134

#24 Activity Theory as a Lens for Examining Teachers' Enactment of the Competency-Based English Teaching Curriculum.....141

Thanh Vinh Chau, Thao Quoc Tran.....141

#26 Enhancing Coherence and Cohesion in IELTS Writing through Generative AI Feedback: An Exploratory Study 155

Nhu Hoang Ngoc Quynh & Sa Tran Si Nguyen 155

#27 Evaluating the Impact of Vocabulary Learning through Online Websites and Digital Resources on University EFL Students: Teachers' Perspectives.....172

Nguyễn Thuỷ Hương 172

#29 English-Major Students' Perceptions of AI Tools for Academic Research: A Quantitative Case Study at a Vietnamese University.....183

Tran Thanh Tu, Tran Nam Anh 183

#34 University Students' Attitude and Intentions Towards The Use of Chatgpt and Canva in a Project-based Learning Activity 201

Vi Le Bao 201

#40 INVESTIGATING ENGLISH-MAJOR UNDERGRADUATES' READINESS FOR AI-SUPPORTED ACADEMIC WRITING: A MULTIDIMENSIONAL PERSPECTIVE.....222

Tran Phuong Anh , & Ngo Thi Thu Hien 222

#41 Students' Perceptions of Google Meet in ESL-Oriented Online Cambridge Exam Preparation at Private Language Centers in Ho Chi Minh City.....240

Ngô Thị Thu Hiền.....240

#42 Examining the Relationship between AI-Supported Self-Regulated Learning and Speaking Self-Efficacy among Vietnamese EFL University Students.....260

Hoang The Hai Nguyen 260

#44 Application of AI in Teaching and Learning Business English at HUFLIT: strengths, weaknesses, and solutions.....289

Nguyen Thi Bach Yen 289

#45 TQM-oriented Service-Learning for EFL-ESL Transition: Social Responsibility at Pham Van Dong University 305

Võ Duy Quân 305

#46 Examining The Factors Affecting The Employment of ChatGPT in Translation Among English Majors at HUFLIT as Vietnamese Education Transitions From EFL To ESL Orientation 315

Huong Le.....315

#49 Developing Intercultural Pragmatic Competence for Real Life Interaction through Role-Play: Teaching the Speech Act of Requests in Vietnamese English Communication Classes in the Transition from EFL-Oriented to ESL-Oriented Pedagogy 370

Nguyen Hoang Vuong Anh, Nguyen Tan Khoa 370

#50 Using Short-Form Videos (TikTok/Reels) to Enhance Real-Life Listening and Speaking Skills for A2 EFL Students: A Pilot Study at Hùng Vương University of Ho Chi Minh City 393

Ông Hồng Thanh.....393

#54 Enhancing English Syntactic Competence through AI-Supported Instruction: A Quasi-Experimental Study of English-Majored Sophomores at HUFLIT.....424

Myphuongnguyen 424

#55 A Literature Review of Appraisal Theory in Language Education: Research Developments and Implications for ESL Teaching 449

Phan Tuan Ly, Hoang Ngoc Thanh Truc[2,3] & Pham Thi Thuy Dung[2,3].....449

#56 Exploring Undergraduate Students' Use of AI for Academic Presentations in Higher Education 468

Nguyễn Thị Hồng Liên, Nguyễn Thị Thanh Vân 468

#58 ASSESSING THE EFFECTIVENESS OF QUIZIZZ IN IMPROVING VOCABULARY LEARNING OF EFL FRESHMEN STUDENTS AT HUFLIT 482

LÊ THY PHƯƠNG THẢO.....482

#59 Self- Reflection in Using GenAI for Learning: Perceptions of English-Majored Students at Binh Duong University- Ca Mau Campus.....494

Liet Thoai Phuong Lan 494

#61 Exploring EFL Freshmen's Motivation for Choosing a Career in Teaching English: A Case Study at A Public University in Ho Chi Minh City 507

Son Le Anh, Tai Phan Tan , Thanh Lam Thanh, Tam Huynh Ngoc, Vy Nguyen Pham Tuong, Hieu Duong Thi Le & Ha Nguyen Thi 507

#63 Vietnamese Students' Readiness for English as a Second Language: The Role of AI-Supported Learning Environments 522

Vu Thanh Loan 522

#64 Bridging the Gap Between Classroom Learning and Real-World Communication in ESL: Curricula and Authentic Assessments 536

Trang Phan Thi Quynh.....536

#65 Aviation English Majors' Perceptions of the Use of AI-based Language Learning Tools for Self-regulated English Speaking Practice.....563

Thị Nguyệt Ánh Nguyễn, Thị Ngọc Quý Nguyễn 563

#67 LEVERAGING GENERATIVE AI TO ADDRESS VIETNAMESE EFL LEARNERS' CHALLENGES IN ENGLISH LITERATURE: A CLOSE READING AND RUBRIC-BASED ASSESSMENT FRAMEWORK FOR THIRD-YEAR UNIVERSITY STUDENTS 577

Nguyễn Thị Thư Hiên 577

#71 INTEGRATING CLASSDOJO INTO WARM-UP ACTIVITIES TO SUPPORT EFL LEARNERS' WRITING PERFORMANCE AT A VIETNAMESE LANGUAGE CENTER.....606

Chi Nguyen, Tran Le Khanh Phuong.....606

#79 EXPLORING THE EFFECTS OF AI-MEDIATED LOW-STAKES PRACTICE ON SPEAKING ANXIETY IN VIETNAMESE EFL CLASSROOMS: A QUASI-EXPERIMENTAL STUDY 622

Minh Hang Le 622

#80 INSTRUCTORS' DIGITAL COMPETENCE AFFECTING STUDENT FEEDBACK AND PERSONALIZED LEARNING IN UNIVERSITY EDUCATION637

Nguyen Thanh Tuan, Nguyen Phuong Bao Tran.....637

#81 THE TEACHER AS A FACILITATOR IN CULTIVATING ENGLISH-RICH HOME ENVIRONMENTS: A PARENTAL COACHING APPROACH FOR PRIMARY EDUCATION 650

Tham Truong Thi , Thanh Phan Thi 650

#82 AI-ASSISTED ESL LEARNING AND STUDENT MOTIVATION: CHALLENGES AND OPPORTUNITIES FOR SOFTWARE ENGINEERING EDUCATION 663

Do Duc Bich Ngan 663

#83 UNPACKING THE BARRIERS: KEY DETERMINANTS OF PUBLIC SPEAKING ANXIETY AMONG ENGLISH MAJOR UNDERGRADUATES AT HUFLIT 681

Vo Quoc Viet 681

#84 ESL INTEGRATION IN JAPANESE READING INSTRUCTION: EXPLORING THE ROLE OF ENGLISH IN SUPPORTING JAPANESE READING COMPREHENSION AMONG SENIOR JAPANESE-MAJOR STUDENTS AT VNU-HCM UNIVERSITY OF SOCIAL SCIENCES AND HUMANITIES 694

Vu Doan Lien Khe 694

#85 FACTORS ASSOCIATED WITH THE SCIENTIFIC RESEARCH COMPETENCE OF HUFLIT ENGLISH LANGUAGE STUDENTS IN PROJECT 1 AND PROJECT 2 707

Quoc Bao Mai 707

#87 EFL STUDENTS' USE OF CHATGPT IN ARGUMENTATIVE ESSAY WRITING: INSIGHTS FROM A QUALITATIVE STUDY 730

Tran Nhu Uyen Le, Diem Bich Huyen Bui 730

#89 EFL TEACHERS' PERCEPTIONS OF AI-SUPPORTED AND DIGITAL TECHNOLOGIES IN IELTS TEACHING FOR HIGH-QUALITY STUDENTS 744

Thai, Ho-Van 744

#92 BEYOND ACCURACY: PRAGMATIC FORCE IN AI-MEDIATED TRANSLATING AND INTERPRETING TRAINING 754

Dương Thị Hoàng Oanh, Nguyễn Xuân Đạt 754

#94 EXPLORING THE INTERPLAY OF MOTIVATION, LEARNER AUTONOMY, AND ACADEMIC PERFORMANCE AMONG ENGLISH-MAJOR STUDENTS AT HUFLIT: A MIXED-METHODS STUDY 771

Ma Kinh Lac Che,, Chau Hoang Anh Phan, Gia Bao Duong & Minh Quyen Nguyen

..... 771

#98 BEYOND THE PROMPT: A METACOGNITIVE ANALYSIS OF STUDENT AI INTEGRATION IN HIGHER EDUCATION 802

Hien Ngoc Nguyen Le 802

#99 AI-POWERED GAME-BASED VOCABULARY PRACTICE FOR ENHANCING VOCABULARY RETENTION AMONG YOUNG VIETNAMESE EFL LEARNERS: A QUASI-EXPERIMENTAL STUDY 825

Trang Trinh Khanh Lien, Chi Ho Phuong Nguyen 825

#100 CREATING ENGLISH-RICH LEARNING ENVIRONMENT BEYOND THE CLASSROOM 842

Pham Thi Bich Tien 842

#105 EXPLORING FIRST-YEAR ENGLISH-MAJORED STUDENTS' PERCEPTIONS OF PEDAGOGICAL TRANSLANGUAGING IN AN EMI CONTEXT 857

Thinh Tien Vu, Cam Diem Nhi Hoang 857

#106 THE USE OF ARTIFICIAL INTELLIGENCE TOOLS IN ESSAY WRITING AMONG FIRST-YEAR ENGLISH MAJORS 875

Nguyen Phuong Thao 875

#107 THE INTRINSIC MOTIVATION OF GENERATION Z AT COFER IN VIETNAM'S SHIFTING ESL CONTEXT 886

Le Nguyen Hoang Nam 886

#110 AI-SUPPORTED WRITING SCAFFOLDING AND PERCEPTIONS OF ACADEMIC INTEGRITY AMONG EFL JUNIORS AT HUFLIT 900

Anh Tuan Nguyen Le, Ngoc Huyen Thai Thi, Minh Chau Diep 900

#113 CHALLENGES AND PEDAGOGICAL RESPONSES IN TRANSITIONING TO ESL-ORIENTED TEACHING: A MIXED-METHODS CLASSROOM CASE STUDY IN A VIETNAMESE UNIVERSITY 921

Huỳnh Thị Thùy Dung[1] 921

#115 CONCEPTUALIZATION OF TEACHERS' RESEARCH MINDSET AS A COMPLEX DYNAMIC SYSTEM IN THE ENGLISH LANGUAGE TEACHING CONTEXT 938

Hạnh Nguyễn Thị Minh, Xuan Mai Le, Trang Le Diem Bui 938

#116 THE EFFICACY AND STUDENT PERCEPTIONS OF AI-GENERATED VERSUS TEACHER FEEDBACK ON UNIVERSITY STUDENTS' L2 WRITING DEVELOPMENT: A MIXED-METHODS STUDY 957

Anh Tran Duc 957

#118 STUDENT PERCEPTIONS OF CHATGPT-GENERATED FEEDBACK IN IELTS SPEAKING PART 3 PRACTICE: A QUALITATIVE STUDY 981

Diem Bich Huyen Bui, Vo Man Nghi Nguyen 981

#120 DIGITAL TRANSFORMATION IN FORMATIVE ASSESSMENT: THE DESIGN, IMPLEMENTATION, AND EVALUATION OF TESTMOZ FOR ENGLISH GENERAL COURSES AT VIETNAM AVIATION ACADEMY 999

Trinh Phan Tu 999

 # 122 ENGLISH TEACHER IDENTITY AND AGENCY IN A VIETNAMESE EDUCATIONAL CONTEXT: RESPONDING TO ESL-ORIENTED EXPECTATIONS 1016

Thanh Hai Ha, Thi Thu Hien Nguyen 1016

#123 THE PERCEIVED USE OF METACOGNITIVE ONLINE READING STRATEGIES AMONG STUDENTS AT AN EMI UNIVERSITY IN VIETNAM ... 1032

Thinh Tien Vu, Thi Thuy Dung Dang 1032

#126 LEARNERS' PERCEPTIONS OF TEACHER ORAL FEEDBACK IN VSTEP SPEAKING PREPARATION: IMPLICATIONS FOR ESL-ORIENTED INSTRUCTION IN VIETNAM 1051

Doan Ngoc Sang, Pham Huu Duc 1051

#127 SHORT-FORM VIDEO PLATFORMS AS DIGITAL TOOLS FOR ESL VOCABULARY DEVELOPMENT: EVIDENCE FROM VIETNAMESE UNIVERSITY STUDENTS 1066

Nguyen Thi Nuong, Phan Thi Thanh Tam, Nguyen Thanh Dung, Duong Thi Thu Uyen

..... 1066

#128 USABILITY TESTING OF WEB-BASED DICTATION TOOLS FOR ESL LISTENING COMPREHENSION 1090

Thao Nguyen Thi Da, Ngoc Phung Tuan 1090

#129 VIETNAM'S ESL TRANSITION: ADAPTING STRUCTURED LITERACY IN MICRO-LEVEL MATERIAL DESIGN FOR PRIMARY CLASSROOMS 1111

Pham Trang Thu, Phan Hoang Thuy Linh 1111

#132 EXPLORING AI SELF-EFFICACY IN EFL STUDENTS: IMPLICATIONS FOR LEARNING PERFORMANCE 1123

Phuong Vo Thi Bich 1123

#134 INTEGRATING CHATGPT INTO EFL WRITING INSTRUCTION: EFFECTS OF THE MEDIATING ROLE OF STUDENTS' PERCEPTIONS AND PERCEIVED ACADEMIC WRITING PERFORMANCE 1140

Nguyễn Thị Thiên Phương, Nguyễn Thị Ngọc Uyển 1140

#135 AN ANALYSIS ON EFL VIETNAMESE SOPHOMORES PRONUNCIATION IN A SPEAKING COURSE 1159

Tran Thi Minh Thu 1159

#136 THE PEDAGOGICAL IMPACT OF ARTIFICIAL INTELLIGENCE TOOLS ON ENGLISH MAJORS' WRITING SKILLS AT HO CHI MINH CITY UNIVERSITY OF FOREIGN LANGUAGES - INFORMATION TECHNOLOGY. 1176

Nguyễn Minh Tuấn 1176

#139 STUDENT EXPECTATIONS OF ENGLISH LANGUAGE TEACHERS: NAVIGATING THE SHIFT TO VIETNAM'S ESL TRANSITION AND AI INTEGRATION 1195

Nguyễn Lê Uyên 1195

#140 RETHINKING ASSESSMENT IN THE SHIFT FROM EFL TO ESL: A CRITICAL REVIEW OF NO-EXAM CLASSROOMS 1214

Thi Thanh Thao Nguyen, Thi Dieu Duyen Le, Thi Hong Tham Chau, Thanh Tinh Bui[4], Van Doan Tran[5] 1214

#142 ENGLISH-MAJOR STUDENTS' READINESS FOR REAL WORLD ENGLISH USE IN AN ESL-ORIENTED CONTEXT: A CASE STUDY AT HUFLIT 1223

Loan Nguyen Thi Thanh 1223

#144 ENGLISH PHARMACEUTICAL TERMINOLOGY AT LHU FACULTY OF PHARMACY: CURRENT USAGE, LEARNING DIFFICULTIES, AND PEDAGOGICAL IMPLICATIONS 1236

Nguyen Huu Thinh, Pham Huu Duc 1236

#146 EXPLORING THE EFFECTS OF SPEAKEASY AI APPLICATION ON SPEAKING FLUENCY AMONG FIRST-YEAR ENGLISH-MAJOR STUDENTS AT HUFLIT 1246

Nguyen Thi Huyen Tran 1246

#149 A FEASIBILITY STUDY OF A GLOBALIZATION AND INTERCULTURAL COMMUNICATION PROGRAM AT THE UNIVERSITY OF LANGUAGES AND INTERNATIONAL STUDIES, VIETNAM NATIONAL UNIVERSITY HANOI: A QUINTUPLE HELIX MODEL 1264

Tran Thanh Nhan, Vu Thi Kim Ngan, Chu Thanh Van 1264

#151 THE USE OF ARTIFICIAL INTELLIGENCE IN ACADEMIC WRITING AND ITS RELATIONSHIP WITH SELF-REGULATED LEARNING: A SYSTEMATIC REVIEW 1290

Thach Huong Giang, Cao Thanh Cong, Trinh Hong Linh 1290

#155 FROM REPRODUCTION TO INTERACTION: EXTENDING ESP LISTENING PRACTICES TOWARD COMMUNICATIVE LISTENING IN AN ESP COURSE AT THE DIPLOMATIC ACADEMY OF VIETNAM (DAV) 1303

Giang Hoang Thi Linh 1303

#156 EXPLORING THE IMPACTS OF OVER-RELIANCE ON GENERATIVE AI TOOLS IN ESSAY WRITING: A CASE STUDY OF HUFLIT SOPHOMORES 1317

Pham Kim Quoc, Le Hoang Yen Vy, Tran Nguyen Thu Ha, Bui Hoang Gia Han, Tran Thi Truc Duyen 1317

#159 ENGLISH LISTENING AND NOTE-TAKING: CHALLENGES AND STRATEGIES FOR HUFLIT ENGLISH MAJORS 1340

Bùi Thị Thu Ngân 1340

#161 BRIDGING THE RHETORICAL GAP: THE CONDITIONAL ROLE OF PERCEIVED DISCURSIVE PROXIMITY IN LEARNING AUTHORIAL VOICE FROM MODEL TEXTS 1358

Tran Huu Thien, Pham Huu Duc 1358

#164 INTEGRATING AI LITERACY AND DELIVERY INSTRUCTION IN PUBLIC SPEAKING CLASSES FOR VIETNAMESE UNDERGRADUATE STUDENTS 1374

Mai Thị Mỹ Hạnh 1374

#172 HUFLIT STUDENTS' PREPARATION FOR GRADUATION CERTIFICATES: PERCEIVED EFFECTIVENESS OF AI-GENERATED LEARNING SUPPORT 1386

Quốc Bảo Mai*, Thị Cẩm Vân Nguyễn[1,2], Đình Bảo Hân Nguyễn[1], & Nguyễn Nhựt Tuyết Hồ[1] 1386

#176 EQUITY, INCLUSION, AND ACCESS IN HIGHER EDUCATION: CHALLENGES FACED BY STUDENTS FROM DIVERSE SOCIOECONOMIC BACKGROUNDS 1409

Trang Nguyen, Vy Nguyen 1409

#178 AI-DRIVEN SELF-DIRECTED LANGUAGE LEARNING: STRATEGY USE AND GENDER PERSPECTIVES AMONG VIETNAMESE EFL UNDERGRADUATES 1426

Huynh Thanh Trung, Chau Thi Kim Ngan, Vo Thi Tien, & Tran Quoc Thao 1426

#180 AI-SUPPORTED PROBLEM-BASED LEARNING FOR ENHANCING UNIVERSITY STUDENTS' PUBLIC SPEAKING PERFORMANCE 1446

Thi Thao Trang Nguyen 1446

#182 STUDENT PERCEPTIONS OF AI-ASSISTED LEARNING AND ITS PERCEIVED ASSOCIATION WITH TEACHER–STUDENT INTERACTION IN EFL SPEAKING CLASSES: AN EXPLORATORY SELF-REPORT STUDY WITH NON-ENGLISH MAJOR UNIVERSITY STUDENTS 1463

Thanh Thuy Nguyen 1463

#184 ENHANCING STUDENT ENGAGEMENT IN INTERPRETATION COURSES THROUGH AI-ASSISTED SELF-PRACTICE 1481

Dang Thi Anh Tuyet 1481

#186 ARTIFICIAL INTELLIGENCE IN ENHANCING LISTENING AND NOTE-TAKING SKILLS: EVIDENCE FROM FISRT-YEAR STUDENTS IN VIETNAM 1498

Tran Tan Phat, Le Ngoc Anh Thu 1498

#189 AI AS A "FACE-NEUTRAL" SPACE: REDUCING ANXIETY AND BOOSTING WILLINGNESS TO COMMUNICATE AMONG VIETNAMESE ENGLISH MAJORS IN THE EFL-TO-ESL TRANSITION 1512

Ngoc Giang Tran, Thi Phuong Mai Do, Thi Cam Trang Thai, Thi Kim Nhung Luu

..... 1512

#190 BETTER WRITING OR LESS THINKING? AI OVER-RELIANCE IN EFL WRITING CLASSROOMS 1535

Thanh Thuy Nguyen 1535

#193 TEACHER BELIEFS, IDENTITY, AND PROFESSIONAL AGENCY IN ESL-ORIENTED TOURISM EDUCATION: A PLS-SEM ANALYSIS IN VIETNAM 1554

Phong Hai Tran[1] 1554

#196 FROM TRADITIONAL TESTING TO ALTERNATIVE ASSESSMENT: EXAMINING THE RELATIONSHIPS AMONG CLASSROOM-BASED ALTERNATIVE ASSESSMENT, STUDENT ENGAGEMENT, AND PERCEIVED COMMUNICATIVE COMPETENCE IN A VIETNAMESE ESL HIGHER-EDUCATION CONTEXT 1577

Phan Tran Duc Lien 1577

#197 HUFLIT ENGLISH MAJORS' PERCEPTIONS OF AI-SUPPORTED TRANSLATION LEARNING IN A VIETNAMESE EFL CONTEXT 1599

Nam Anh Tran, Thanh Tu Tran 1599

#198 INTEGRATION OF LOCAL HISTORY AND CULTURAL MATERIAL INTO ENGLISH LANGUAGE TEACHING AS BASIS FOR THE DEVELOPMENT OF MULTICULTURAL PERSONALITY 1619

Gulnur Polkina, Evgeniya Tsyganova, Svetlana Radionova 1619

#201 INTEGRATING DEBATE-BASED PEDAGOGY TO SUPPORT VIETNAM'S TRANSITION FROM EFL TO ESL 1630

Nguyen Thuy Hong Van, Guillermo Reyna 1630

INTRODUCTION

The TESOL International Conference 2026 (TIC 2026), organized by INTESOL Vietnam, VEDUCON, and Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), continues to provide a forum for educators, researchers, and practitioners to share ideas and experiences in English language education. TIC 2026 takes place at a time when English teaching in Vietnam is changing rapidly. English is increasingly used not only as a subject of study but also as a means of communication in education, work, technology, and international exchange. This shift encourages us to rethink the traditional EFL classroom and explore more communicative and ESL-oriented approaches that connect classroom learning with real-life language use.

Artificial intelligence is another important part of this changing landscape. AI tools are creating new possibilities for teaching, learning, feedback, assessment, and independent practice. At the same time, they raise important questions about over-reliance, academic integrity, critical thinking, and the changing role of teachers. TIC 2026 provides a space to discuss these opportunities and concerns from both research and classroom perspectives.

The papers in this volume reflect these developments. They cover a wide range of topics, including AI-supported language learning, writing and speaking, translation and interpreting, learner autonomy, digital competence, assessment, teacher development, communicative teaching, authentic language use, and the transition from EFL to more ESL-oriented learning environments. What brings these studies together is a shared concern with how English can be taught and learned more effectively in a rapidly changing world. The contributions also remind us that technology alone cannot transform education. Meaningful change still depends on good pedagogy, thoughtful teachers, active learners, and learning environments that encourage genuine communication.

The Editorial Board sincerely thanks all authors, reviewers, presenters, participants, organizers, and partner institutions for their contributions to TIC 2026. We hope that these proceedings will provide useful ideas for further research, classroom practice, and professional collaboration in Vietnam and beyond. We welcome comments and feedback from researchers, educators, and readers at **tesol@huflit.edu.vn**.

EDITORIAL BOARD OF THE PROCEEDINGS

#1 The Impacts of Various Teaching Methods and Assessments Applied in Business Communication Course

Thái Mỹ Linh[1]

ABSTRACT

With globalization increasing the demand for proficient business communication, it is crucial to equip ESL learners with the skills required for professional environments. Business Communication is a compulsory subject in the Business English program at the Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology. Based on student-centered approaches, task-based language teaching, and ongoing assessment, this study investigates how interactive teaching methods and these assessment strategies support students' communication proficiency, engagement, and readiness for workplace communication. The study examines students' perceptions of the course and the impacts of role-playing, case studies, task-based activities, presentations, projects, quizzes, peer assessment, and AI-enhanced learning tools. Data were collected from 155 student questionnaires and classroom observation checklists. The findings show that students strongly valued the course, with 95.4% confirming its importance for communication skill development. Interactive methods, case studies, hands-on activities, and presentations received highly positive responses, while project-based assessment and peer assessment were also perceived as useful, although they required careful management to reduce stress and improve feedback quality. Practically, the findings suggest that Business Communication courses should combine real-world tasks, structured peer feedback, and balanced assessment schedules to enhance both communication competence and learner confidence.

Keywords: Business Communication, teaching methods, evaluations, impacts

1. Introduction

Business Communication is a compulsory subject for juniors majoring in Business English at HUFLIT. This course is important because they are provided with useful knowledge such as how to communicate in the digital- age workplace, how to plan and participate in face-to-face and virtual meetings, and the techniques to use and develop expert writing. They also have the chance to practice and improve practical skills like professionalism and business etiquette skills, intercultural workplace skills. However, this course is high-pressure because they have to absorb considerable knowledge and acquire professional skills in business settings, which they almost have no experience and background knowledge before. In addition, with the immense amount of theory combined with several tasks to put the theory into practice, students feel overloaded if appropriate methods are not applied. According to the results of the studies carried out by Tajeddin (2018) and Barnett-Itzhaki (2023), a variety of innovative teaching methods, including case studies, presentations, projects and the applications of digital tools, can boost students' learning outcomes and level of engagement in all activities.

Another important thing that needs to be taken into consideration is how to design tests that can accurately and reliably evaluate students' competencies. To have an accurate evaluation of students' abilities, Asemanyi (2015) showed the importance of using a variety of assessment techniques, such as formative assessments, peer reviews, and assignments.

More recent studies also support the need for learner-centered and task-oriented approaches in higher education. Goodwin (2024) argues that student-centered teaching

[1] *Ho Chi Minh University of Foreign Languages - Information Technology, Vietnam; Email: linhtm@huflit.edu.vn*

promotes engagement, autonomy, and critical thinking, while recent work on task-based language teaching shows that authentic, goal-oriented tasks can improve learner motivation, communicative competence, and classroom participation (Huang, 2024; Mudinillah et al., 2024; Waluyo & Wangdi, 2024). These findings are particularly relevant to Business Communication because students are expected not only to understand business concepts but also to perform realistic communication tasks such as meetings, interviews, negotiations, presentations, and professional writing. Therefore, the present study responds to the practical problem of how to design teaching and assessment activities that reduce passive learning, increase participation, and prepare students for communication demands in future workplaces.

From the challenges mentioned above, the author applied various methods of teaching as well as testing to keep students motivated and help them achieve their goals.

2. Literature Review

The field of business communication education has come up with remarkable innovations including the integration of diverse teaching approaches and evaluation methods. These innovations have occurred thanks to the continuous development of technology, which contributes greatly to enhance students' involvement, facilitate their learning process, and equip themselves for the professional career path in the future.

❖ Teaching Methods and Assessment Strategies

Various teaching methods have been applied to improve the effectiveness of these courses. Interactive lectures, projects, presentations, case studies, etc., are among the most outstanding approaches. Research by Berry (2013) highlights the impacts of technology on teaching business communication, emphasizing the need for strategies that align with a culturally diverse audience. The use of social media platforms and collaborative tools has been shown to create a more dynamic and participatory learning environment, fostering more engagement and practical skill development. Additionally, Çelik and Alpan (2023) found that innovative teaching methods can improve certain aspects of student development.

Assessment methods by to role in evaluating and improving students' proficiency. To provide accurate evaluations and meaningful feedback, effective assessment tools must uphold the principles of fairness, validity, and reliability. Furthermore, incorporating a variety of assessment approaches enhances the learning experience by offering students multiple opportunities to showcase their skills. This not only accurately evaluates students' performance but also encourages their continuous improvement and effort to make progress.

➢ Project-Based Learning

The concept of learning through projects is not new and is increasingly applied in educational institutions around the world. It is undeniable that projects play a vital role in assessing and developing students' teamwork, problem- solving, and interpersonal skills. By replicating real working environments with work-related issues, these projects provide students with opportunities to engage in group work to fulfil the tasks. Asemanyi (2015) noted that group projects help students tackle more complex problems, assign roles and responsibilities, and share diverse viewpoint- skills that are essential in the workplace. Additionally, group projects have been shown to enhance students' time management skills, critical thinking abilities, and deepen their understanding through discussions. According to Yuliansyah & Ayu (2021), in language teaching, project-based tasks can yield tangible products such as presentations or creative products. Solomon (2003) demonstrated that project-based learning requires learners to work together to find solutions to challenging problems that are realistic

and based on the curriculum. Accordingly, the value of project-based tasks lies not only in the final product but also in the process of achieving the final goal (Kerimkulova et al., 2016).

- **Task-based language teaching**

According to Larsen-Freeman and Anderson (2011), tasks have a clear meaning and purpose, and learners need to communicate in a foreign language to achieve the goal. Therefore, this method can be considered to encourage learners to use a foreign language both in the classroom and in real-life contexts. Teachers must set clear goals to accelerate learning by engaging learners in various tasks with clear outcomes. This method focuses on a learner-centered approach. In addition, in lesson design, according to Anderson (2017), each learner's needs should be considered or other needs that may arise during the process of receiving the assigned task. Therefore, not only learning goals but also learning opportunities can be created.

In this study, task-based language teaching was operationalized through job-search tasks, mock interviews, negotiation role-plays, professional email writing, CV and cover-letter preparation, and case-based problem-solving activities. These tasks were later reflected in the results through students' responses to interactive teaching methods, hands-on activities, presentations, and project-based assessments. Therefore, the TBLT framework is not treated as a separate theory but is embedded in the course activities that required students to use English to complete realistic business communication tasks.

- **Simulations**

Simulations are forms of active learning experiences and interactive activities that create realistic scenarios, allowing students to apply language skills in a controlled environment (Lean et al., 2006). Research by Johnson et al. (2018) demonstrates that role-playing activities improve learners' communication skills, empathy, and adaptability. Simulations, such as virtual sales environments, allow students to practice without real-world consequences.

- **Peer-reviewing tasks**

Research has demonstrated that peer-reviewing helps students develop motivation and confidence in the learning process, thereby significantly improving learning outcomes. Research by Hojeij and Baroudi (2018) examined the impact of peer-reviewing on student motivation and engagement. The results showed that learners expressed positive attitudes towards peer feedback. Furthermore, students who participated in peer feedback received several benefits such as self-reflective thinking skills, problem solving and critical thinking, development of learner autonomy, promotion of self-regulation in learning and development of social skills (Hansen & Liu, 2002). Involving students in the assessment process through peer evaluations enhances their critical thinking skills and provides diverse feedback. This method encourages student responsibility and active participation in learning. According to Boon (2018), peer assessments offer varied feedback and promote critical evaluation skills, fostering student accountability and active learning. Peer evaluations have been shown to improve students' understanding of success criteria, increase their engagement in the learning process, and develop their interpersonal skills. Additionally, peer evaluations create a positive classroom environment where students can assess their own and their peers' work, leading to better learning outcomes.

- **Presentations:**

In educational contexts, oral presentations are now a common way to assess a variety of student competencies. Both verbal and nonverbal communication skills are essential for success in the workplace, and this approach is especially good at evaluating them. Nadolski et

al. (2021) highlight the value of oral presentations as an assessment tool for developing and evaluating students' presentation competence. Their study found that a well-designed formative assessment method using analytic rubrics significantly improved students' oral presentation performance. Oral presentations require students to organize and deliver messages effectively while demonstrating skills such as clarity, structure, vocal delivery, and non-verbal communication. These skills are particularly relevant to professional communication, where individuals need to present information clearly and persuasively to an audience.

- **Case Studies:**

By analyzing and presenting case studies, students can bridge the gap between theoretical knowledge and real-world business scenarios. This method not only assesses their oral and written communication skills but also enhances their critical thinking and problem-solving abilities. Asemanyi (2015) highlighted that case studies are an effective way to connect academic concepts with practical application. Through this approach, students improve their ability to tackle real business challenges, thereby honing their critical thinking and problem-solving skills. Additionally, presenting their analysis and solutions clearly and persuasively helps evaluate their communication skills. Case studies provide a practical learning experience, equipping students with tactical knowledge and skills that are directly applicable to their future careers.

- **Innovative methods applied during the course**

The student-centered approaches were applied during the course. Students were required to work in groups of 5 to 6, each of which was responsible for a project to design a training program to improve intercultural communication within an international company. The project needs to cover sessions on the significance of cultural identities, communication styles, etiquettes and taboos, cultural norms and values and conflict resolution. Each group would choose a country to focus on providing deep knowledge with illustrations like clips, short movies, etc.,

Task-based activities were designed to provide students with opportunities to practice. Students were asked to search LinkedIn to get information about job vacancies and choose the one that is most suitable for them. Then they had to conduct and participate in mock interviews. Based on the theoretical knowledge about low-context and high-context culture, they had to apply the principles of communication styles in convincing the interviewers. They also engaged in role-playing exercises to practice negotiation skills, focusing on persuasive communication, active listening, and finding mutually beneficial solutions. They were also assigned with various writing exercises: complete their own CV and cover letter, draft professional emails for various business situations.

The third task is presentations. Students prepared and delivered presentations on hot business-related topics like how businesses can use social media platforms to reach their potential customers and react in the center of a scandal.

Additionally, related to case studies, students were required to analyze a real-life business crisis. They needed to read a lot and choose one case that impressed them in terms of its impact. They had to evaluate the company's communication strategy, identify what was done well, and suggest improvements. Another topic for case study is about the success or failure of a company's customer service communication. Students analyzed the communication methods used, the results, and suggested the best practices. One last option is investigating a successful or unsuccessful marketing campaign. Students discuss the communication strategies employed, the target audience, and the campaign's impact.

After the course, the researcher would like to evaluate the significance of the course from the students' point of view as well as the impacts of various teaching methods and evaluations. Three questions were raised:

Research Questions

1. Is it really necessary to include Business Communication in the curriculum?
2. How do different teaching methods impact student engagement and learning outcomes in business communication courses?
3. What are the effects of various assessment strategies on the development of business communication skills?

3. Methods

Pedagogical Setting & Participants

The survey was conducted in five Business Communication classes that the researcher oversaw teaching in the first semester of the school year 2024-2025. The learners were third-year students majoring in Business English, Faculty of Foreign Languages, HUFLIT. In previous courses, students had studied compulsory subjects such as Business English, Office Communication, Negotiation Skills, and Introduction to Management. These subjects equipped students with specialized vocabulary, knowledge and skills in reading and understanding business-related documents, drafting correspondence in business transactions, presenting on specialized topics, etc. 155 out of 208 students in five classes completed the questionnaire.

Design of the Study

A mixed-method strategy was used by the researcher. The questionnaire was designed including five scales, ranging from strongly agree to strongly disagree, from 1 to 5 respectively. It consisted of three sections: the significance of the course, the impacts of various teaching methods and the students' feedback on the necessity of various evaluation forms. It was designed in Google form and sent to students to the system of Moodle of the university. The researcher asked the students to complete the questionnaire to collect quantitative information. The qualitative component was conducted through a structured classroom observation checklist used weekly during the course. The checklist focused on observable indicators such as participation level, willingness to speak, teamwork, confidence in presentations, response to peer feedback, anxiety during assessments, and ability to complete communicative tasks. After each class, the researcher recorded short field notes beside each checklist item. In addition, the researcher also carried out classroom observations to have a better understanding of how students react to the innovated methods. During the course, students' attitudes, behaviors, and participation in each activity were noted down in the observation checklist weekly.

Data collection & analysis

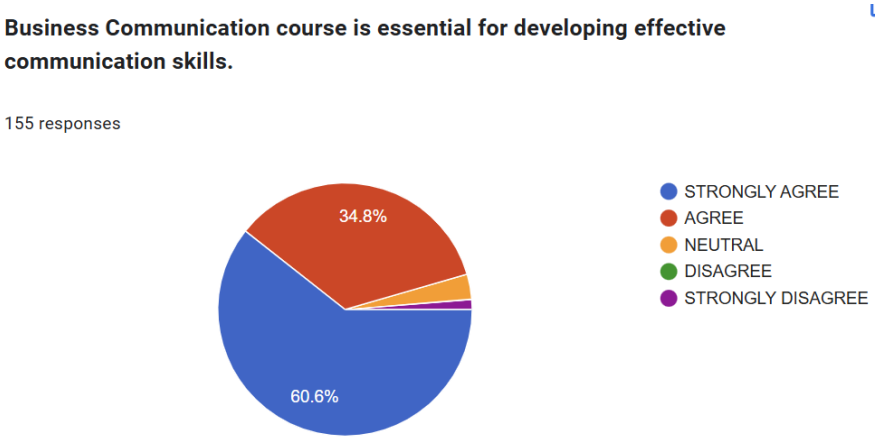
The researcher created the questionnaire on Google Forms and sent it on Moodle in the last week of the semester. This was to ensure that students had sufficient time and experience to evaluate the course and provide reliable feedback. They also did not need to worry about the lecturer's reaction when giving negative responses. The researcher had one week to collect the data. Quantitative survey data were summarized by frequency and percentage to identify students' overall perceptions of each teaching method and assessment form. Qualitative observation notes were analyzed through a simple thematic coding process. First, the researcher reviewed weekly notes several times to become familiar with recurring patterns. Second, similar observations were coded, such as "active participation," "improved confidence,"

"assessment anxiety," "teamwork difficulty," and "quality of peer feedback." Third, these codes were grouped into broader themes that corresponded to the research questions: course significance, teaching-method impact, and assessment impact. The qualitative themes were then used to explain and support the quantitative survey results.

4. Results/Findings

Figure 1

The importance of Business Communication courses in developing effective communication skills.



An overwhelming majority of students acknowledged the importance of this course, as seen in Figure 1. 34.8% of students agreed and 60.6% strongly agreed that the course is essential to improve their communication abilities. This means that an overall number 95.4% of the students appreciated the course's beneficial influence on their competence for effective communication.

Only a tiny proportion of students had neutral viewpoints. Compared to the overwhelming positive responses, this group was not significant. Only two out of 155 students—or just 1.3%—strongly disagreed with the course's importance.

The theoretical knowledge and practical skills that the students gained from this course are obviously advantageous. The overwhelming positive feedback asserts that the course is greatly beneficial to improving students' communication skills.

Figure 2

Interactive teaching methods (group discussions, role-playing) enhance students' engagement.

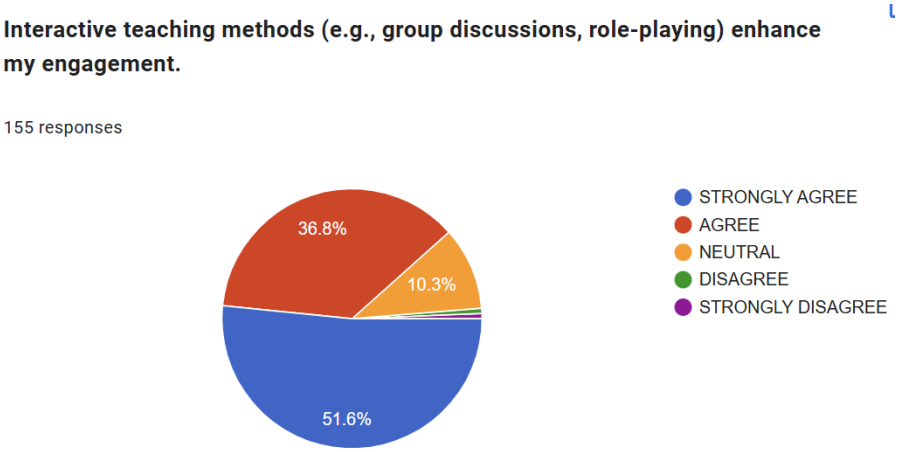


Figure 2 illustrates how various interactive teaching techniques can contribute to increase students' engagement in the learning process. According to the results, approximately 90% of respondents had extremely positive feedback on the application of diverse learner-centered methods used in the course. This significant proportion shows that students highly valued the effectiveness that these interactive approaches- including role-playing, group debates, and other collaborative tasks- bring to help them improve.

A tiny percentage of students, only 1 out of 155, or 0.6%, strongly disagreed with the effectiveness of these approaches. Additionally, the same percentage of students disagreed, emphasizing the positive comments outweigh the negative ones. The data highlights the effectiveness of applying interactive teaching methods.

In addition, the notes in the researcher's observation checklists during the 12-week period provide supportive evidence for these conclusions. It was noticeable throughout this period that the great majority of students engaged actively in all the interactive tasks and activities that were organized every week in class. The high degree of participation shows that students were not only interested in these approaches but also appreciated that these methods were truly helpful for their education and future career. Thanks to the application of these methods, a more dynamic and captivating learning environment was established. As a result, students were more engaged in the learning process, which improved their motivation in the process of exploring knowledge.

Figure 3

The roles of case studies and real-world examples

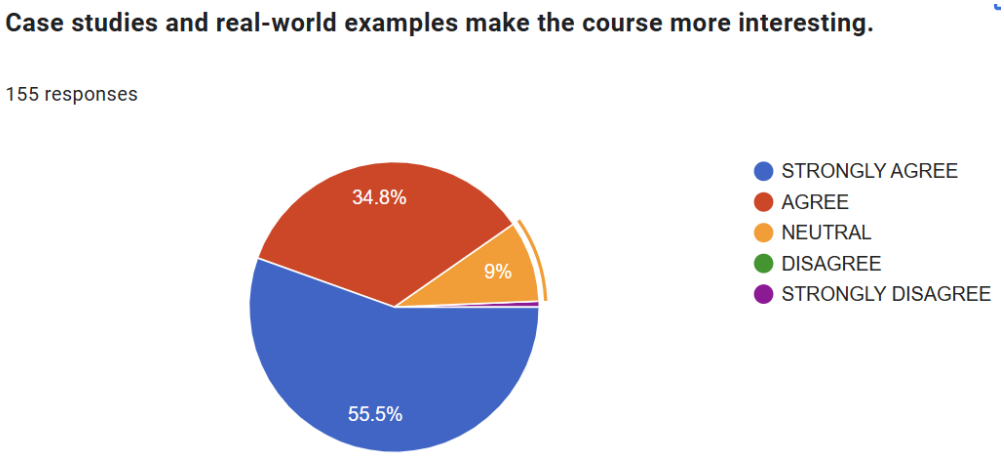


Figure 3 demonstrates the importance of case studies and real-world examples in drawing and maintaining students' attention during the course. According to the data, an overwhelming majority of respondents felt satisfied with the analysis of case studies and business-related examples. 34.8% of students agreed and 55.5% strongly agreed that this was very successful at maintaining their interest. In other words, 90.3% of students had positive feedback about the use of case studies and real-world examples in the curriculum.

Additionally, according to the notes from the observation checklists, the researcher realized that when students participated in discussions and debates based on case studies, they became very enthusiastic and interested. These tasks promoted critical thinking and active engagement. Students were able to apply theory to real-world situations, which motivated them in the process of exploring knowledge.

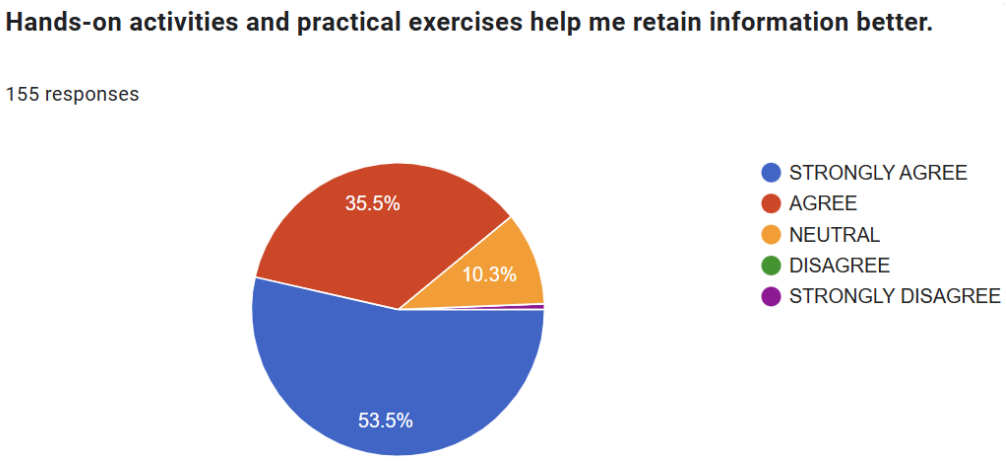
Only 9% of students expressed no opinion about how successful case studies and real-world examples are. This proportion can be implied that these methods could be improved and modified further to better cater to the diverse linguistic backgrounds and proficiency levels of every learner. A tiny minority, 0.6% of students (1 out of 155) strongly disagreed that case studies and real-world examples were helpful, which is not remarkable.

The results show that case studies and real-world examples play a crucial role in drawing in and retaining students' attention. The overwhelmingly positive responses demonstrate how effective these methods are in creating an engaging and motivating learning environment.

To sum up, findings from Figure 3 and notes from the observation checklists clearly show that using case studies and real-world examples is a very good way to grab students' interest and improve their involvement in the course. In addition to making learning more fun, these techniques greatly enhance critical thinking and problem-solving abilities, which are crucial for students' success in university and the future career paths as well.

Figure 4

The roles of hands-on activities and practical exercises in helping students retain information



It is evident from Figure 4 that a sizable majority of students-89% acknowledged the value of practical exercises and hands-on activities in aiding in their memory retention. Because the course gives students a thorough understanding of business topics, covering a wide range of concepts, theories, principles, and more, it would be very difficult for students to remember and apply the knowledge they have learned if proper teaching techniques were not used. These exercises and hands on activities offer students the opportunity to engage with theoretical concepts more practically, enhancing their ability to retain and apply knowledge in real-life situations. This strategy greatly differs from conventional lecture-based techniques, which make students passively learn new knowledge and as a result hinder them from remembering new business concepts. Furthermore, only a small minority were unaware of the usefulness of these techniques. While the neutral responses indicate some variation in learning preferences, the overall positive impact of the different approaches is clear.

The results of on-going exams and assessments provide more evidence of the efficacy of various techniques. Based on their results in these assessments, it can be concluded that students' comprehension ability and memory of the subject matter have clearly been strengthened by the frequent inclusion of these exercises and hands-on activities. In summary,

the result and evidence from on-going evaluations highlight the significant role of practical exercises and hand-on activities in the learning process. These methods not only enhance students' ability to retain information but also prepare them to apply it in real world situations.

Figure 5

The application of presentations

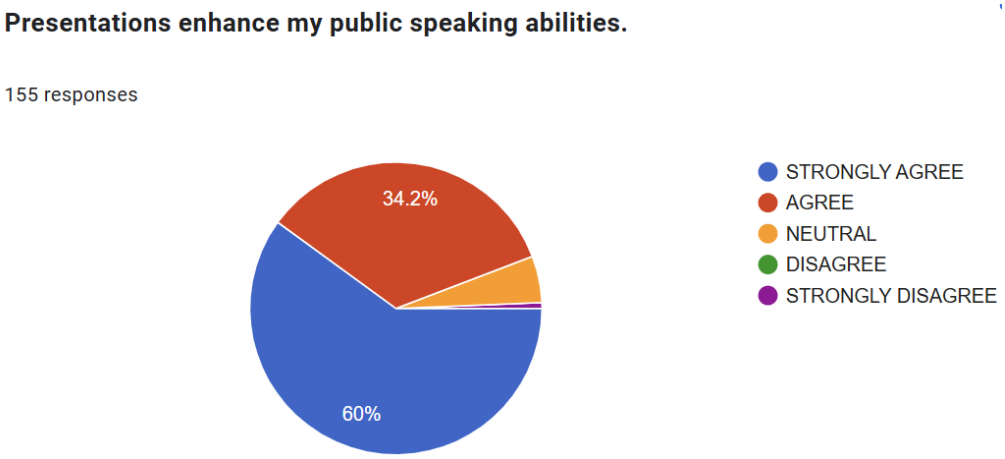


Figure 5 demonstrates how delivering presentations can enhance students' confidence and their ability to speak in public. Presentations clearly play a crucial role in building their self- esteem, as reflected by 34.2% of students who agreed and the 60% strongly agreed that presentations greatly improve this challenging skill. This indicates that an overwhelming number of students overall acknowledged the beneficial effects of presentations on their ability to express ideas in public. These conclusions are further supported by the observations recorded in the checklists. According to the notes, about 70% of students expressed their extreme anxiety when they had to give a presentation for the first time. This anxiety showed out in several ways, including stuttering, repeating sentences or even phrases several times, and completely avoiding eye contact with the audience. Furthermore, a lot of students found it difficult to engage with their audience, which is a typical problem for those who are exposed to public speaking the first time.

But as the course went on, there was a remarkable change. Students became more and more confident in their next presentations. They began to communicate more smoothly, interact with their audience more effectively, and maintain eye contact. The continuous practice and feedback they received from their lecturers directly contributed to their increased professionalism and confidence as presenters. Additionally, only 5.2% of students expressed neutral views regarding the effectiveness of presentations. It can be explained that these students could not overcome the tension and stress when delivering a speech and 12 weeks seems not enough for them to struggle.

Just 0.6% of students disagreed that presentations were effective. The positive feedback outnumbered the negative ones, which demonstrates the importance of presentations in helping students become more confident and proficient speakers.

In summary, the results shown in Figure 5 and the observational data clearly indicate that giving presentations is an effective way for boosting students' confidence and communication skills in public. Students gain more experience and practise, their initial discomfort gradually diminished, leading to improved performance and greater professionalism. The widespread agreement among students on the benefits of presentations

underscores their importance in the curriculum and establishes them as a key tool for developing essential communication skills.

Figure 6

Project-based assessments

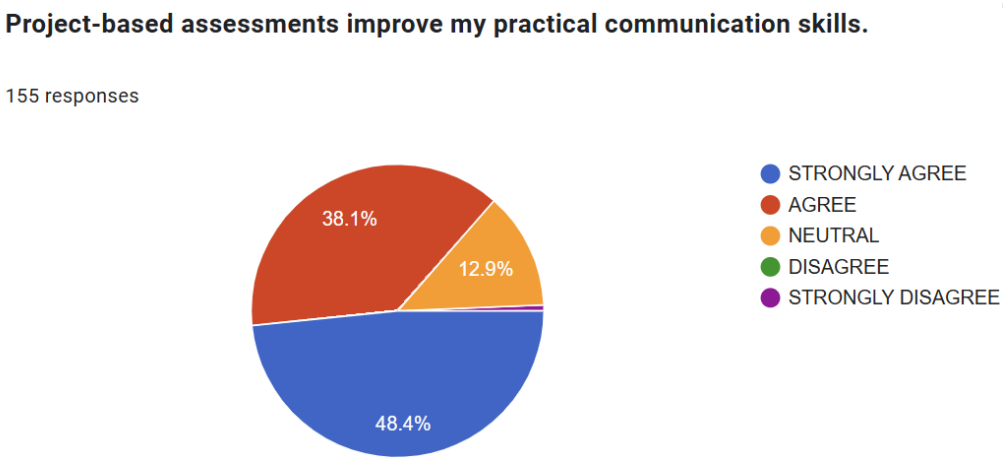


Figure 6 shows students' feedback about project-based assessments. 38.1% of the students agreed, and nearly half strongly agreed, that project-based evaluations greatly improve their communication abilities. This indicates that an overwhelming proportion 86.5% of students in all acknowledged the importance of project-based evaluations in improving their capacity for clear communication. It can be concluded that project-based learning is an effective method for students at this stage of academic career.

12.9% of students expressed no opinion about the benefits of project-based assessments in helping them improve their skill. Based on the notes in the observation checklists, the researcher realized that some students experienced stress when working on projects for a variety of reasons. The fact that group members had different schedules can be challenging to find a time fitting everyone. Moreover, unavoidable conflicts arose, impeded the project's advancement and increased their stress levels. Another important reason is that some students might not have good time management skills and interpersonal skills needed to successfully complete group projects. These skills are essential for completing the projects.

Despite these challenges, it can be said that the project-based approach is still the best option for this subject at the highest level of university education. Instead of relying on the lecturer, students are required to take part in projects and familiarize themselves with the tasks to be fulfilled. By doing this, they are prepared for their future professional employment. This approach gets them ready for the realities of the corporate world, where projects are becoming more essential.

In conclusion, project-based evaluations have a lot of benefits in terms of skill development and preparation for real-world business. Through encouraging self-reliance, teamwork, and clear communication, project-based learning gives students the skills they need to be successful in university and in the workplace.

Figure 7

The effectiveness of frequent quizzes and tests

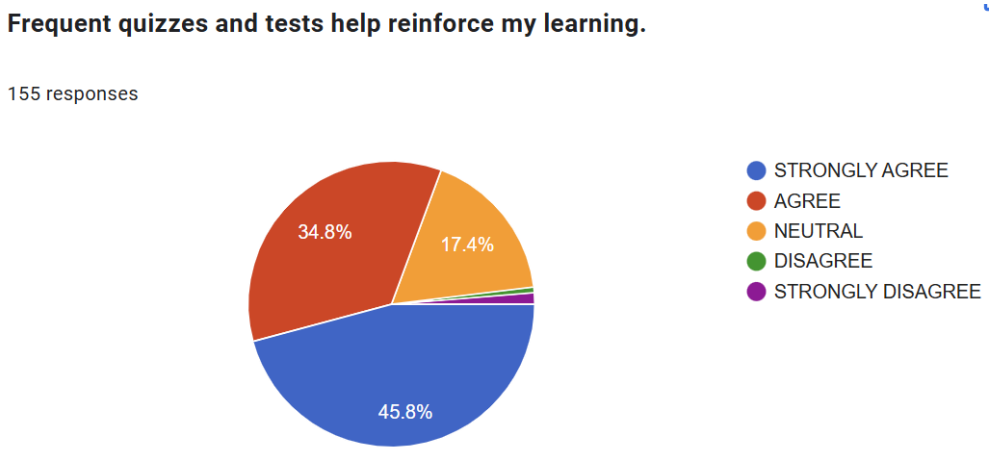


Figure 7 shows the students' opinions about the implementation of frequent quizzes and tests in class. According to the results, 45.8% and 34.8% of students respectively strongly agreed and agreed that regular tests and quizzes aid in their learning. They acknowledged the advantages of frequent evaluations. According to the notes from the researcher's checklist, when students received ongoing feedback on their progress from these constant assessments, they knew which areas needed improving. As a result, there will be fewer mistakes or weaknesses in the next tests.

It is noticeable that a comparatively high proportion 17.4% of students had no idea about the benefits of frequent tests and quizzes. According to the notes from the checklists, some students confirmed that they experienced tension and anxiety when they had to take tests and quizzes on a regular basis. This feedback is important because it emphasizes how frequent assessments might affect students in a negative and unexpected way.

According to the requirements in the course detailed outline, at least three mini-tests must be taken throughout a 12-week period. The purpose of this frequent evaluation is to ensure that students can consistently track their progress and identify their weakness. However, this puts a lot of strain on the students and the lecturer as well. The pressure to perform well on these tests led to increased stress and anxiety. This has an adverse effect on the students' learning process. Students must continuously study for and do well on these tests, and lecturers must constantly create and grade these tests. To tackle the issue, teachers must consider the different forms of these tests to reduce the pressure on students.

In conclusion, although the data shows that most students recognize the benefits of regular tests and quizzes, the comments about stress and the sizeable percentage of ambivalent answers emphasize the necessity for a flexible strategy. Teachers should try to create tests that can effectively evaluate the students' skills and reduce their stress as well.

Figure 8

The role of peer assessment in helping students improve

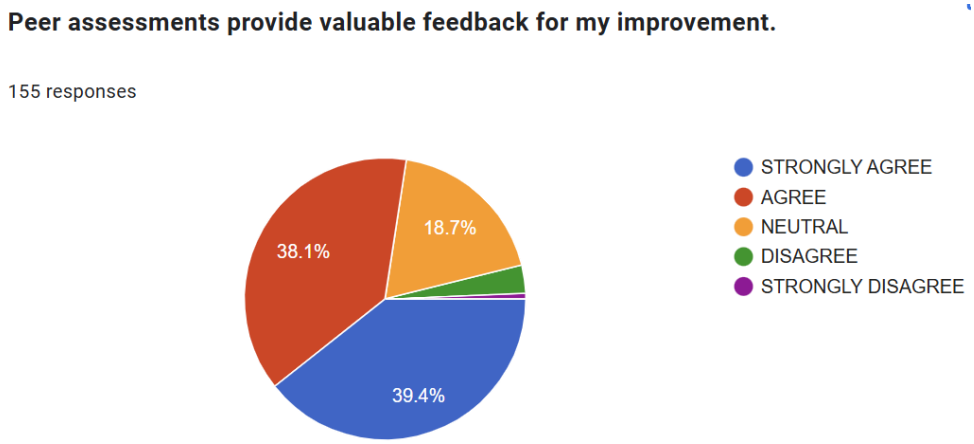


Figure 8 illustrates students' opinions about their peers' feedback in helping them improve their skills.

According to the results, the percentages of respondents who strongly agreed and agreed that peer assessments much enhance their improvement were approximately the same, accounting for 39.4% and 38.1%, respectively. Noticeably, 18.7% of respondents had neutral ideas about how helpful peer evaluations are. The highly subjective remarks and criticism that some students got on their friends' work can be interpreted in conjunction with the observation checklist notes. Despite of the fact that many of the remarks were helpful and constructive, some feedback was not objective, based on prejudice and emotional judgement which might have made some students not value their peer assessments. These results are supported by the notes from the observation checklists, which show that constructive and justified feedback contributes greatly to the improvement as well as the motivation of the students. They frankly showed that they did not appreciate or value their friends' comments if it is opinionated or prejudiced. Therefore, it is crucial to guide students on how to provide their peers with unbiased, helpful, and constructive feedback. It is advisable to provide them with a range of criteria to evaluate their peers in various tasks. Only 3.2% of students didn't agree that peer evaluations are valuable for them in developing their skills.

Generally, positive responses about peer assessments outnumbered the negative ones, which highlights the potential of this as a valuable teaching tool. Peer assessments promote collaborative learning and a deeper understanding of the theory by encouraging students to critically analyze one another's work. Peer assessments strengthen students' ability to evaluate the work of others and simultaneously they learn a lot from their peers thanks to the observation process.

In summary, the results and data from the observational notes show that peer evaluations are crucial for assisting students in developing their knowledge and abilities. The overwhelmingly positive responses indicate that peer assessments are an indispensable part of the learning process. To improve peer assessments' efficacy and guarantee that every student gains from this collaborative learning strategy, it is necessary to tackle the problems with feedback quality.

5. Discussion

According to this research, most students- 95.4% positive responses all together- agreed strongly or agreed that the course plays a vital role in improving their communication skills, which are essential for becoming employable in the competitive job market. A study by Ismail

(2011) concluded that the prospect of being employed is higher amongst graduates with a good competency of English. As a result, the students' feedback shows that being equipped with the skills and knowledge in this course, they are more confident that they will improve their employability. Another important finding is that approximately 90% of students gave favorable comments, claiming that interactive learner-centered teaching techniques like role-playing and group discussions greatly increase their involvement in all activities, which makes learning more dynamic and efficient. The results are not out of the expectation of the author and supported by the data from the checklist notes. The learner-centered methods proved to be extremely effective in gaining students' attention and interest. In addition, most students found case studies and real-world examples to be very engaging. The same number of students agreed that practical exercises and hands-on activities are crucial for long-term retention of business knowledge because they allow theoretical concepts to be applied practically. In brief, innovated approaches that focus on the students' needs are highly appreciated and contribute to increasing their level of satisfaction.

Despite of the fact that project-based tasks might be stressful, 88.1% of students acknowledged the benefits they bring. Based on the notes from the observation checklists, the researcher confirmed that students improved several skills during the process of doing the projects. They had to analyze real-world problems and brainstorm to discuss innovative solutions. Therefore, they can be more creative and motivated as well to fulfill the tasks. These projects also greatly foster teamwork and improve their interpersonal skills. They had the chances to explore the issues in depth, leading to better retention and understanding. In short, project-based evaluations are the most suitable option for senior students because they are prepared for future work.

It cannot be denied that frequent quizzes and tests create stress for students. Surprisingly, 80.6% of students felt that they were quite beneficial. It can be explained that the variety of test forms like mini tests, quizzes, presentations, projects, etc. to evaluate different skills reduces the boredom of doing tests and students recognize the benefits of these constant evaluations.

Finally, the number of students who found peer assessments to be useful was not so high, accounting for 77.5%. More than 22% of students were unhappy with their peers' comments. The writer thinks that because there were some biases in some peer evaluations and comments, nearly a quarter of students were unsatisfied.

This study supports earlier research that highlights the value of hands-on, interactive learner-centered teaching strategies. It asserts the idea that active learning techniques, such project-based learning and practical exercises, improve student interest and memory. In line with past research emphasizing the importance of communication skills in the labor market, the results support the importance of these abilities in enhancing employability. Furthermore, the study's findings about the value of peer evaluations and presentations support other studies that found similar techniques to be essential for fostering students' self-assurance and teamwork. In conclusion, the study adds to the increasing amount of data demonstrating creative and student-centered teaching strategies in higher education.

6. Conclusion

The findings from this study highlight the importance of this course to the students' employability and adaptability in the future workplace. In addition, a variety of student-centered teaching methods and innovative forms of on-going assessments have been proved to enhance their interest and motivation in learning, help them take ownership of their learning

and acquire important skills to meet the challenging demands of the labor market. From the students' perspective, these various teaching methods and evaluations, including role-playing, case studies, projects and presentations play a significant role in assisting them improve their communication proficiency and explore the knowledge in depth.

In direct response to the three research questions, the findings confirm that Business Communication is perceived as necessary in the curriculum because it supports students' employability, confidence, and readiness for workplace communication. The findings also show that learner-centered teaching methods, especially role-playing, case studies, task-based activities, presentations, and practical exercises, increased students' engagement and helped them connect theory with real-life business contexts. Finally, the results indicate that assessment strategies such as projects, quizzes, presentations, and peer assessment contributed to skill development when they were clearly structured, balanced in workload, and supported by constructive feedback. These conclusions demonstrate a clear alignment between the research questions, survey findings, and classroom observation themes.

One thing that needs to be taken into consideration is that although regular examinations and quizzes are good for consolidating knowledge, they should be balanced to prevent unnecessary stress for students. During 12-week semester, 3 compulsory tests and a project seem to be a burden for students because they must meet deadlines in other courses. I suggest that there should be some modification in the course detail outline so that tests are not stressful for both the test maker and test takers.

Another important thing is that even though peer evaluations are a good solution for teachers as they save time and for students as they get valuable feedback from their peers, they must be carefully managed to guarantee helpful and constructive criticism. From my point of view, to achieve this goal, the teacher should provide students with a guideline in which there are specific requirements and criteria to evaluate their peers' performance.

Despite offering insightful information, the study had a few drawbacks. Only 155 students took part in the survey, which might not accurately reflect all the students majoring in Business English.

A larger and more varied sample size should be considered in future studies to improve the findings' generalizability and reliability. Furthermore, investigating how well these teaching strategies work in more classes would support and broaden the conclusions.

7. Acknowledgments

My sincere appreciation goes out to the following people and institutions for their vital contributions to this study:

The Department of Business English, for supporting me to carry out this research. Their assistance was essential in making the research efforts possible.

The students who took part in the study, whose eager participation and candid criticism were crucial to its success. My results and conclusions were based on their contributions.

For their helpful critiques and support throughout the several phases of this study, I am grateful to my peers and colleagues. Their suggestions enabled me to increase the study's overall quality and hone my methodology.

Thank you to family and friends for your steadfast understanding and support during the study process. Their patience and support were crucial.

I would like to express my sincere gratitude to everyone who helped with this study, and I recognize that without their combined efforts, it would not have been possible.

References

Asemanyi, A. A. (2015). An Assessment of Students' Performance in Communication Skills: A Case Study of the University of Education Winneba. *Journal of Education and Practice*, 6(35), https://files.eric.ed.gov/fulltext/EJ1086368.pdf

Barnett-Itzhaki, Z., Beimel, D., & Tsoury, A. (2023). Using a variety of interactive learning methods to improve learning effectiveness: insights from AI models based on teaching surveys. Online Learning, 27(3). https://doi.org/10.24059/olj.v27i3.3575

Berry, P. (2012). The impact of globalization and technology on teaching business communication: reframing and enlarging world view, methods, and content. American Journal of Business Education (AJBE), 6(1), 57-66. https://doi.org/10.19030/ajbe.v6i1.7483

Boon, S. (2018). Using peer assessment as an effective learning strategy in the classroom. *Perspective Article,* https://my.chartered.college/impact_article/using-peer-assessment-as-an-effective-learning-strategy-in-the-classroom/

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Çelik, Ö. C., & Alpan, G. (2023). The Impact of an Effective Communication Course with Enhanced Student Engagement on Communication Skills and Empathic Tendency of Preservice Teachers. *Educational Process International Journal,* 12(2), 33-58.

Farashahi, M., & Tajeddin, M. (2018). Effectiveness of teaching methods in business education: a comparison study on the learning outcomes of lectures, case studies and simulations. The International Journal of Management Education, 16(1), 131-142. https://doi.org/10.1016/j.ijme.2018.01.003

Goodwin, J. R. (2024). What's the difference? a comparison of student-centered teaching methods. Education Sciences. https://doi.org/10.3390/educsci14070736

Huang, X. (2024). Research on the effectiveness of task-based language teaching in higher education. *Adult and Higher Education, 6*(1), 77–82. https://doi.org/10.23977/aduhe.2024.060111

Khamis, N., & Wahi, W. (2021). Enhancing business communication skills: an action research. International Journal of Asian Social Science, 11(7), 355-364. https://doi.org/10.18488/journal.1.2021.117.355.364

Mudinillah, A., Rahmi, S. N., & Taro, N. (2024). Task-based language teaching: A systematic review of research and applications. *Lingeduca: Journal of Language and Education Studies, 3*(2), 102–115. https://doi.org/10.70177/lingeduca.v3i2.1352

Nadolski, R. J., Hummel, H. G. K., Rusman, E., & Ackermans, K. (2021). Rubric formats for the formative assessment of oral presentation skills acquisition in secondary education. Educational Technology Research and Development, 69, 2663–2682. https://doi.org/10.1007/s11423-021-10030-7

Waluyo, B., & Wangdi, T. (2024). Task-based language teaching in academic english higher education: a case of low-proficiency learners in thailand. 3L The Southeast Asian Journal of English Language Studies, 30(3), 246-264. https://doi.org/10.17576/3l-2024-3003-17

Biodata

The researcher is a full- time lecturer in Ho Chi Minh city University of Foreign languages and Information technology. She has nearly 30 years' experience teaching Business English. She's interested in methodology, language testing and related fields. She regularly attends conferences to learn from other researchers.

#2 EVALUATING PROGRAM LEARNING OUTCOME CHIEVEMENT THROUGH A DATA-INFORMED OUTCOME-BASED EDUCATION APPROACH: EVIDENCE FROM ENGLISH-MAJORED UNDERGRADUATES AT HUFLIT UNIVERSITY

Tran Trung Hieu[1] , Tran Kim Tuyen[2]

Abstract

Outcome-Based Education (OBE) has become a dominant framework for curriculum design and quality assurance in higher education. Within this framework, program learning outcomes (PLOs) serve as key indicators of students' expected competencies upon graduation. However, limited empirical evidence exists regarding how specific PLO domains contribute to students' employability readiness, particularly in English language studies programs in Vietnam. This study investigated students' perceptions of the achievement of four program learning outcomes—English language proficiency and competency, majoring-specialization competencies, digital literacy, and second foreign language competence (SFL), and examined their relationships with employability readiness. A quantitative survey was administered to 261 students enrolled in the English language studies program at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). Data were analyzed using descriptive statistics, reliability analysis, Pearson correlation, one-way ANOVA, and multiple regression analysis. The findings revealed that students generally reported positive perceptions of all PLOs and employability readiness (ER). All scales demonstrated satisfactory reliability, with Cronbach's alpha coefficients ranging from .883 to .957. Correlation analysis showed significant positive relationships between all PLO domains and employability readiness. Multiple regression analysis revealed that the four PLO domains jointly explained 59.2% of the variance in employability readiness ($R^2 = .592$, $F = 92.995$, $p < .001$). Hierarchical regression further demonstrated that the PLO variables contributed an additional 54.2% of explained variance beyond demographic and academic characteristics ($\Delta R^2 = .542$, $p < .001$), highlighting the substantial role of program learning outcomes in shaping students' ER.

Keywords: Outcome-Based Education (OBE), program learning outcomes (PLO), English Language Studies programs, employability readiness (ER), Common European Framework of Reference for Languages (CEFR)

[1] *Lecturer, Faculty of Foreign Languages, Hochiminh City University of Foreign Languages and Information Technology, Hochiminh Ciy, Vietnam; Email: hieutt1@huflit.edu.vn*

[2] *Lecturer, Faculty of Foreign Languages, Hochiminh City University of Foreign Languages and Information Technology, Hochiminh Ciy, Vietnam; Email: tuyentk@huflit.edu.vn*

1. INTRODUCTION

1.1 Background of the study

In recent years, higher education institutions (HEI) have faced increasing pressure to prepare graduates who can meet the demands of a rapidly changing and highly competitive labor market. Beyond disciplinary knowledge, employers increasingly value graduates who possess strong communication skills, digital competence, adaptability, and professional readiness. Consequently, graduate employability has become an important indicator of higher education quality and effectiveness.

To address these expectations, many universities have adopted Outcome-Based Education (OBE), an educational approach that emphasizes the achievement of clearly defined learning outcomes. Within OBE, Program Learning Outcomes (PLOs) serve as key benchmarks for curriculum design, teaching, assessment, and quality assurance. The effectiveness of an academic program is therefore often evaluated based on the extent to which students achieve the intended learning outcomes (LOs).

In Vietnam, OBE has been widely implemented as part of ongoing higher education reforms and accreditation requirements. English Language Studies programs, in particular, are expected to equip students with a combination of language proficiency, disciplinary knowledge, digital literacy, and multilingual competencies that support their future careers. While these competencies are commonly embedded in program learning outcomes, empirical evidence regarding their contribution to employability readiness (ER) remains limited.

1.2 Research problem

Although previous studies have examined language competence, curriculum quality, and graduate employability separately, relatively few have investigated how different PLO domains collectively influence students' ER. This gap highlights the need for further research on the relationship between PLOs and ER within English language education programs.

Within this context, the English Language Studies program at Hochiminh City University of Foreign Languages – Information Technology (HUFLIT), Vietnam, provides an appropriate case for investigating outcome-based curriculum effectiveness. The program is designed according to OBE principles and incorporates multiple PLO domains, including English language proficiency and competency (PLO3), majoring-specialization competencies (PLO4), digital literacy (PLO5), and SFL competence (PLO6). While these outcomes are formally embedded in the curriculum, there is limited empirical evidence on how students perceive their attainment of these outcomes and how such perceptions relate to their overall ER.

Therefore, this study adopts a data-informed OBE approach to examine students' perceived attainment of the four key PLOs and their relationship with the ER.

1.3 Research objectives

This study aims to:

1. Examine students' perceptions of the achievement of the four key PLOs including PLO3, PLO4, PLO5 and PLO6.
2. Investigate the relationships between these four PLOs and the ER.
3. Identify the extent to which PLOs predict the ER.

1.4 Research questions

1. Research question 1: How do English-majored students perceive their attainment of the four PLOs in the English Language Studies program at HUFLIT?
2. Research question 2: Which PLO domains are perceived as the most and least developed by students?
3. Research question 3: To what extent do different PLO domains contribute to students' perceived ER?

1.5 Significance of the study

This study contributes to the growing body of researches on OBE by providing empirical evidence regarding the effectiveness of PLOs in fostering ER. The findings may assist curriculum developers, academic leaders, and quality assurance practitioners in improving educational programs and enhancing graduate outcomes.

2. Literature Review

2.1 Outcome-based education in higher education

OBE has become one of the most influential educational paradigms in higher education worldwide. According to Spady (1994), OBE emphasizes the achievement of clearly articulated learning outcomes rather than the completion of prescribed instructional content. Under this approach, curriculum design, teaching strategies, assessment methods, and quality assurance mechanisms are systematically aligned with the intended graduate competencies.

The growing adoption of OBE reflects broader shifts toward accountability, transparency, and evidence-based quality assurance in higher education. Universities are increasingly required to demonstrate that graduates possess competencies relevant to contemporary workplace demands and societal expectations. Consequently, program effectiveness is evaluated not only through academic performance but also through measurable learning outcomes and graduate success indicators.

Recent scholarship suggests that OBE contributes positively to curriculum coherence, learner-centered instruction, and continuous quality improvement. According to Biggs and Tang (2022), outcome-oriented curriculum design promotes constructive alignment among learning outcomes, teaching activities, and assessment practices, thereby enhancing students' learning experiences. Similarly, Harden (2007) argues that OBE facilitates systematic curriculum evaluation and supports institutions in meeting accreditation standards.

In Southeast Asia, OBE has become a key component of higher education reform. Universities have increasingly integrated outcome-based principles into curriculum development and assessment systems to align educational provision with international quality assurance frameworks. However, despite widespread implementation, empirical studies examining how specific PLOs contribute to graduate readiness remain relatively limited, particularly in language education programs.

2.2 Program learning outcomes and constructive alignment

PLOs constitute the cornerstone of OBE implementation. They define the knowledge, skills, attitudes, and competencies that students are expected to demonstrate upon graduation. Effective PLOs serve multiple functions, including curriculum planning, instructional alignment, assessment design, and stakeholder communication.

The theoretical foundation underlying PLO implementation is Constructive Alignment Theory (Biggs & Tang, 2022). This framework posits that meaningful learning occurs when

intended learning outcomes, teaching activities, and assessment methods are aligned. Such alignment enables students to engage in learning activities that directly support the achievement of desired competencies.

Research consistently demonstrates positive relationships between curriculum alignment and student achievement. Programs with clearly articulated and systematically assessed learning outcomes tend to report stronger academic performance, higher student satisfaction, and improved graduate outcomes. Furthermore, outcome alignment supports institutional quality assurance processes by providing measurable indicators of educational effectiveness.

Within English language education, PLOs typically encompass multiple competency domains, including linguistic proficiency, disciplinary knowledge, intercultural communication, digital literacy, and professional skills. Because language graduates increasingly work in diverse and technology-mediated environments, outcome frameworks must extend beyond language competence alone.

Additionally, according to Kennedy (2007), learning outcomes describe what learners are expected to know, understand, and be able to demonstrate upon successful completion of a programme of study. Clearly articulated learning outcomes therefore serve as the foundation for curriculum design, teaching, assessment, and quality assurance processes.

2.3 English language education, CEFR, and competency development

The Common European Framework of Reference for Languages (CEFR) has become the dominant international framework for language proficiency assessment and curriculum design. Since its introduction, CEFR has provided a common reference system for describing language ability across educational contexts.

In Vietnam, CEFR-aligned learning outcomes have been widely adopted in higher education institutions, particularly within English language programs. Universities increasingly utilize CEFR benchmarks to establish graduation requirements and evaluate language competence. This alignment enhances transparency, comparability, and international recognition of language qualifications.

Nevertheless, contemporary English language education extends beyond linguistic competence. The evolving nature of global communication and workplace practices requires graduates to possess a broader range of professional competencies. Scholars have argued that successful language graduates must integrate language proficiency with disciplinary expertise, intercultural competence, technological literacy, and lifelong learning capabilities.

Digital transformation has further expanded expectations for language professionals. The emergence of artificial intelligence, online communication platforms, and digital collaboration tools has increased the importance of technological competence within language-related professions. Consequently, many English language programs have incorporated digital literacy as a core learning outcome alongside traditional linguistic objectives.

Recent studies indicate that graduates possessing both strong language skills and digital competencies demonstrate greater adaptability and employability in contemporary labor markets. This suggests that language proficiency alone may be insufficient to ensure successful career outcomes, highlighting the need to examine the combined influence of multiple competency domains.

2.4 Graduate employability and employability readiness

Graduate employability has emerged as a central concern for higher education institutions globally. As universities face increasing pressure to demonstrate societal relevance, employability outcomes have become important indicators of educational quality and program effectiveness.

Yorke (2006) conceptualizes employability as a set of achievements, including knowledge, skills, understandings, and personal attributes that increase graduates' likelihood of obtaining employment and succeeding in their chosen professions. More recent frameworks extend this definition by emphasizing adaptability, resilience, lifelong learning, and technological competence.

The literature identifies several dimensions consistently associated with employability, comprising of professional knowledge and disciplinary expertise, communication and interpersonal skills, critical thinking and problem-solving abilities, digital literacy and technological competence, adaptability and lifelong learning capacity, and multilingual and intercultural communication skills.

For graduates of English language programs, ER represents a multidimensional construct. Employers increasingly seek graduates who can communicate effectively in English while simultaneously demonstrating professional competence, technological proficiency, and cross-cultural communication skills.

Previous studies have reported positive relationships between language proficiency and employability outcomes. However, recent evidence suggests that employability is influenced by a combination of competencies rather than a single skill domain. Graduates who possess integrated skill profiles tend to demonstrate greater workplace readiness and career adaptability.

2.5 Empirical studies and research gap

A growing body of researches has examined the effectiveness of OBE implementation in higher education. Existing studies generally report positive impacts on curriculum quality, student achievement, and learning outcome attainment. However, several limitations remain evident in the literature.

First, many studies focus primarily on evaluating individual learning outcomes rather than examining the combined contribution of multiple PLO domains. Second, researches in language education frequently emphasize language proficiency while overlooking broader competencies such as digital literacy and multilingual capability. Third, empirical investigations linking OBE-based learning outcomes directly to employability readiness remain relatively scarce, particularly within Southeast Asian contexts.

Within Vietnam, researches on outcome-based curriculum evaluation have increased in recent years. Nevertheless, studies that simultaneously investigate language proficiency, disciplinary competence, digital literacy, second foreign language competence, and employability readiness are still limited. Consequently, there is insufficient evidence regarding which competency domains contribute most strongly to students' perceived ER.

To address these gaps, the present study examines four key PLO domains: language proficiency and competency (PLO3), majoring-specialization competencies (PLO4), digital literacy (PLO5), and second foreign language competence (PLO6) and investigates their relationships with the ER among English language studies students. By integrating outcome-based assessment and employability perspectives, this study contributes empirical evidence to

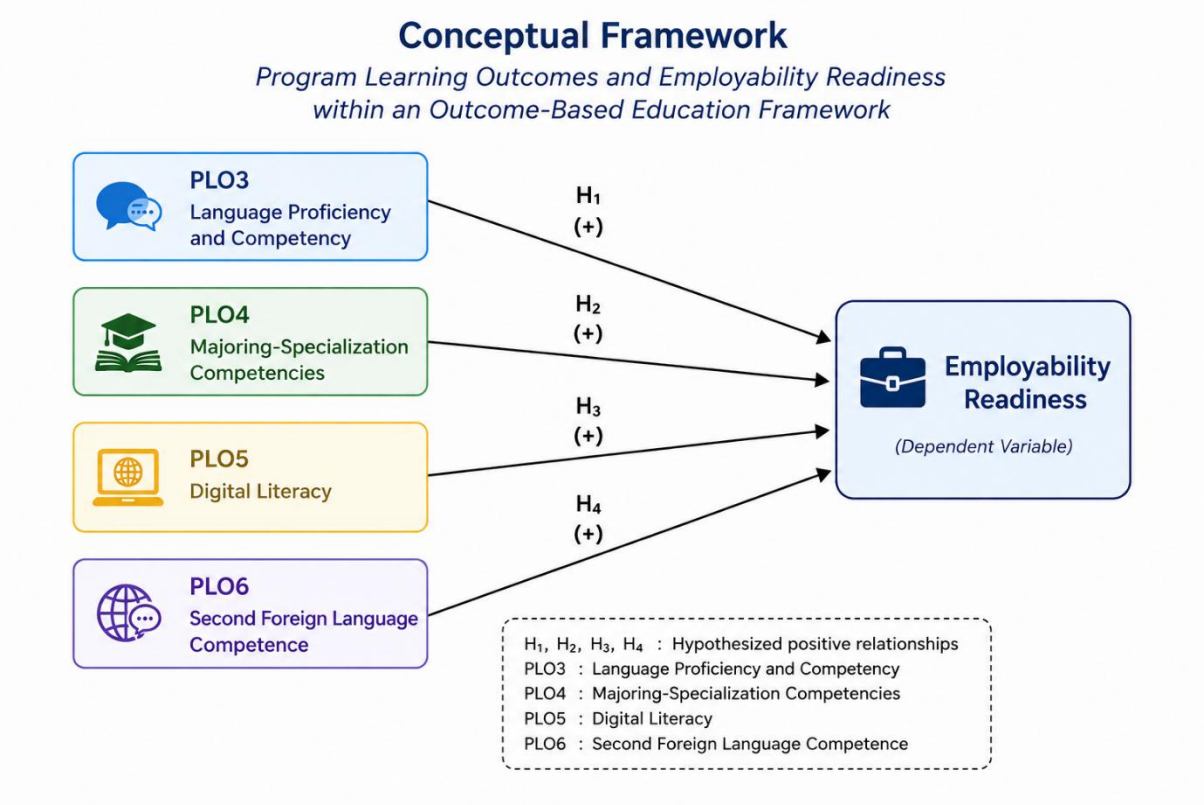
support data-informed curriculum evaluation and quality assurance practices in higher education.

2.6 Conceptual framework

Based on the literature reviewed, the study proposes the following conceptual framework:

Figure

Research conceptual framework



3. METHODOLOGY

3.1 Research Design

This study employed a quantitative correlational research design to investigate the relationships between the four designated PLOs and the employability readiness among English Language Studies students at HUFLIT University.

The study was grounded in the Outcome-Based Education (OBE) framework, which emphasizes the achievement of clearly defined learning outcomes and their contribution to graduate competence. Specifically, the study examined four major PLO domains:

PLO3: English language proficiency and competency;

PLO4: majoring-specialization competencies;

PLO5: digital literacy;

PLO6: second foreign language competence.

ER: employability readiness.

ER was treated as the dependent variable representing students' perceived preparedness for future employment.

A cross-sectional survey design was adopted because it allowed researchers to collect data from a large number of respondents within a relatively short period and examine statistical relationships among variables.

3.2 Research context

The study was conducted within the English Language Studies program at HUFLIT University, Vietnam, a private higher education institution that offers undergraduate and postgraduate programs in languages, business, information technology, and related disciplines.

In response to national higher education reforms and quality assurance requirements, the English Language Studies program which is revised and edited every two years has adopted an OBE framework to ensure the alignment among curriculum design, instructional practices, assessment activities, and intended PLOs. Within this framework, student learning achievement is systematically evaluated to support continuous program improvement and accreditation processes.

The program specifies four key competency domains that are expected to contribute to graduates' professional development and employability: PLO3, PLO4, PLO5, and PLO6. These competencies reflect both disciplinary requirements and the transferable skills increasingly demanded by contemporary labor markets.

As part of the program's quality assurance and curriculum evaluation efforts, regular assessments are conducted to examine students' perceptions of learning outcome attainment. Against this background, the present study investigates the extent to which students perceive achievement of the specified PLOs and examines the predictive relationships between PLO attainment and ER. The findings are expected to provide empirical evidence for curriculum enhancement, outcome-based program evaluation, and evidence-informed decision-making within the English Language Studies program.

3.3 Participants

A total of 261 undergraduate students participated in the study. Participants represented different academic years within the English Language Studies program. Data were collected through voluntary participation using an online questionnaire (Google.docs).

The sample which was chosen on the convenient sampling included students from various specialization majors, including Translation and Interpretation, English Language Teaching, English for Business, English for Offices, English for Logistics and English–Chinese Bilingual Studies.

The participation was anonymous and no personally identifiable information was collected.

3.4 Research instrument

3.4.1 Questionnaire design

The questionnaire items were developed based on the official PLOs of the English Language Studies program and informed by previous literature on graduate employability, outcome-based education, and competency assessment (Yorke, 2006; Tomlinson, 2017; Kennedy, 2007).

The instrument consisted of five constructs:

Construct	Description	Items
PLO3	English language proficiency and competence	5
PLO4	Majoring-specialization competencies	5
PLO5	Digital literacy	5
PLO6	Second foreign language competence	5
ER	Perceived readiness for future employment	6

The final questionnaire contained 26 items.

Responses were measured using a five-point Likert scale: 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree. Higher scores indicated stronger perceptions of outcome attainment and ER.

3.4.2 Pilot testing and questionnaire validation

Prior to the main data collection, the questionnaire was piloted with 40 English Language Studies students who were not included in the final sample. The purpose of the pilot study was to evaluate item clarity, wording appropriateness, and preliminary reliability of the instrument.

Participants were invited to provide feedback regarding item ambiguity and comprehensibility. Based on their comments, several items were revised to improve wording clarity and reduce redundancy.

The pilot data also yielded satisfactory internal consistency coefficients across all constructs (Cronbach's alpha values ranging from .78 to .89), indicating acceptable reliability for subsequent administration.

3.4.3 Reliability analysis

Prior to conducting the main statistical analyses, the reliability of all measurement scales was examined using Cronbach’s alpha coefficients for assuring the internal consistency. Cronbach’s alpha coefficients were interpreted based on the recommendations of Nunnally and Bernstein (1994), who suggested that reliability values of .70 or above indicate acceptable internal consistency.

Furthermore, corrected item-total correlations were also examined for suggesting that all questionnaire items contributed meaningfully to their respective constructs.

3.5 Data collection procedure

Data were collected during the second semester of the academic year through an online survey platform.

Before participation, respondents were informed about the purpose of the study, voluntary participation, confidentiality of responses, academic use of collected data. Students completed the questionnaire anonymously.

After data screening, 261 valid responses were retained for statistical analysis.

3.6 Data analysis

The collected data were analyzed using IBM SPSS Statistics.

Several statistical procedures were employed.

3.6.1 Descriptive statistics

Means (M) and standard deviations (SD) were calculated to examine students' perceptions of learning outcome attainment and ER.

3.6.2 Pearson correlation analysis

Pearson Product-Moment Correlation was employed to examine the relationships among PLO3, PLO4, PLO5, PLO6, and ER.

The assumptions of multiple regression, including multicollinearity and independence of errors, were assessed following the guidelines proposed by Hair et al. (2022).

3.6.3 Multiple regression analysis

Multiple linear regression was conducted to determine the extent to which PLO attainment predicts ER.

The regression model was specified as:

$ER=\beta0+\beta1(PLO3)+\beta2(PLO4)+\beta3(PLO5)+\beta4(PLO6)+\varepsilon$; where:

ER served as the dependent variable; PLO3, PLO4, PLO5, and PLO6 served as independent variables.

Model fit was additionally evaluated using R^2, Adjusted R^2, F-statistic, and Standardized beta coefficients.

3.6.4 Assumption testing

Several assumptions of regression analysis were examined, including Multicollinearity.

Multicollinearity was assessed through Tolerance values, Variance Inflation Factor (VIF).

For Independence of Errors, the Durbin–Watson statistic was used to assess residual independence.

3.6.5 One-Way ANOVA

One-way Analysis of Variance (ANOVA) was performed to investigate whether perceptions of PLO attainment and ER differed across academic year groups in which academic year served as the grouping variable.

Homogeneity of variance was examined using Levene's Test.

When variance homogeneity assumptions were violated or group sizes were unequal, Welch's robust test was consulted to confirm the robustness of ANOVA findings.

Post-hoc analyses using Bonferroni and Tamhane procedures were conducted where appropriate.

3.7 Ethical considerations

The study adhered to standard ethical principles for educational research.

Participants were informed about the purpose of the study. They participated voluntarily; they could withdraw at any time and remain anonymous throughout the research process.

The collected data were used solely for academic research and program evaluation purposes.

4. RESULTS

4.1 Reliability Analysis

Table 1

Reliability Statistics of the Measurement Scales

Construct	Items	Cronbach's Alpha
PLO3	5	.883
PLO4	5	.748 – .932*
PLO5	5	.910
PLO6	5	.957
ER	6	.922

PLO3 achieved a Cronbach's alpha coefficient of .883. For PLO4, Cronbach's alpha coefficients across the specialization-specific sub-majors ranged from .748 to .932, indicating acceptable to excellent reliability. PLO5 demonstrated a Cronbach's alpha of .910, while ER produced an alpha coefficient of .922. Among all constructs, PLO6 demonstrated the highest level of internal consistency ($\alpha = .957$).

Furthermore, all corrected item–total correlations exceeded the recommended threshold of .30, indicating satisfactory item discrimination. Examination of the "Cronbach's Alpha if Item Deleted" values also revealed that removing any individual item would not substantially improve scale reliability. Consequently, all items were retained for subsequent analyses.

Overall, all reliability coefficients exceeded the commonly accepted benchmark of .70 (Hair et al. 2022), confirming that the questionnaire possessed adequate internal consistency for assessing students' perceived attainment of PLOs and ER.

*PLO4 was evaluated across six majoring-specialization groups and consistently demonstrated acceptable-to-excellent reliability.

4.2 Descriptive statistics of program learning outcome attainment

Table 2

Descriptive Statistics of Program Learning Outcome Attainment and Employability Readiness (N = 261)

Construct	Number of Items	Mean Range	SD Range
PLO3	5	3.32 – 3.82	0.824 – 1.046
PLO4*	5	3.30 – 4.50	0.488 – 1.254
PLO5	5	4.15 – 4.22	0.709 – 0.791
PLO6	5	3.46 – 3.76	1.083–1.188
ER	6	3.72 – 4.04	0.746–0.887

The results revealed generally positive perceptions across all constructs. Among the investigated domains, PLO5 received the highest ratings. Item means ranged from 4.15 to 4.22, indicating that students perceived themselves as highly competent in utilizing digital technologies and information resources for academic and professional purposes.

PLO3 received moderately high ratings, with mean scores ranging from 3.32 to 3.82. The highest-rated item was PLO3_2 (M = 3.82, SD = .824), whereas PLO3_5 recorded the lowest mean score (M = 3.32, SD = 1.046), suggesting some variation in students' perceptions across different language-related competencies.

For PLO4, mean scores varied across specialization groups. Relatively high ratings were observed among students in the English Teaching and Business English tracks with several items exceeding a mean value of 4.00, while lower ratings tended to appear in several smaller specialization groups. However, these differences should be interpreted cautiously because of substantial variations in subgroup sample sizes.

PLO6 received comparatively lower ratings than the other PLO domains, ranging from 3.46 to 3.76. In addition, this construct exhibited the largest standard deviations (SDs ranging from 1.08 to 1.19), indicating considerable variability in students' perceived second foreign language competence.

Regarding ER, students reported positive perceptions across all six indicators. Mean scores ranged from 3.72 to 4.04, suggesting that respondents generally felt prepared to enter the labor market upon graduation.

Taken together, the descriptive findings indicate that students perceived relatively high levels of PLO attainment and ER. Among the investigated domains, PLO5 emerged as the strongest area, whereas PLO6 represented a comparatively weaker domain and may warrant further curricular attention.

4.3 Correlation analysis

Table 3

Pearson Correlation Matrix of Program Learning Outcomes and Employability Readiness (N = 261)

Variables	1	2	3	4	5
1. PLO3	—				
2. PLO4	.696**	—			
3. PLO5	.493**	.565**	—		
4. PLO6	.449**	.401**	.375**	—	
5. ER	.633**	.716**	.582**	.444**	—

As shown in Table 3, all correlations were positive and statistically significant at the .01 level.

ER was positively correlated with PLO3 (r = .633), PLO4 (r = .716), PLO5 (r = .582), and PLO6 (r = .444). Among the investigated predictors, PLO4 demonstrated the strongest association with ER, suggesting that majoring-specialization competencies may play a particularly important role in shaping students' ER.

Intercorrelations among the predictor variables ranged from .375 to .696, indicating moderate associations among the constructs. Since all correlation coefficients were below the recommended threshold of .80 (Hair et al., 2022), preliminary evidence suggested that multicollinearity was unlikely to be a concern. This assumption was further confirmed in the

regression analysis, where VIF values ranged from 1.41 to 2.34, well below the recommended cutoff value of 5.0.

While all four PLO domains were significantly correlated with employability readiness, subsequent regression analysis revealed that only certain competencies remained significant predictors when the shared variance among PLOs was controlled.

4.4 Multiple Regression Analysis

Table 4

Multiple Regression Analysis Predicting Employability Readiness from Program Learning Outcomes (N = 261)

Predictor	B	SE B	B	T	P	Tolerance	VIF
Constant	0.482	0.265	—	1.816	.071	—	—
PLO3	0.187	0.069	.212	2.704	.008	.428	2.335
PLO4	0.335	0.081	.325	4.156	< .001	.431	2.320
PLO5	0.257	0.070	.225	3.663	< .001	.697	1.434
PLO6	0.109	0.043	.152	2.504	.013	.711	1.406

Model Summary: $R = .770$; $R^2 = .592$; Adjusted $R^2 = .586$; $F(4, 256) = 92.995$, $p < .001$; Durbin–Watson = 1.953.

Note. ER = Employability Readiness. β = standardized coefficient.

To determine the extent to which PLO predicted ER, a multiple linear regression analysis was conducted.

The overall regression model was statistically significant, $F(4, 256) = 92.995$, $p < .001$. The model produced an R value of .770 and explained 59.2% of the variance in Employability Readiness ($R^2 = .592$; Adjusted $R^2 = .586$), indicating substantial explanatory power.

The Durbin–Watson statistic was 1.953, suggesting that the assumption of independent errors was satisfied.

As presented in Table 4, all four PLO domains significantly predicted ER. Among the predictors, PLO4 emerged as the strongest predictor ($\beta = .325$, $p < .001$), followed by PLO5 ($\beta = .225$, $p < .001$), PLO3 ($\beta = .212$, $p = .008$), and PLO6 ($\beta = .152$, $p = .013$). These findings indicate that higher levels of perceived PLO attainment were associated with greater ER among English-major students.

Furthermore, collinearity diagnostics indicated no evidence of multicollinearity. Tolerance values ranged from .428 to .711, while Variance Inflation Factor (VIF) values ranged from 1.406 to 2.335, all well below the recommended threshold of 5.0 (Hair et al., 2022). Therefore, the regression coefficients can be interpreted with confidence.

Overall, the findings suggest that the four PLO domains collectively represent important determinants of students' ER, with specialization-related competencies exerting the strongest influence.

5. DISCUSSION

5.1 Students' perceived attainment of Program Learning Outcomes

The findings indicate that English-major students generally perceived themselves as having achieved the intended PLOs at satisfactory levels. Across all competency domains, mean scores exceeded the midpoint of the five-point scale, suggesting positive evaluations of learning outcome attainment. These findings provide preliminary evidence that the OBE framework implemented within the English Language Studies program has been effective in supporting students' academic, professional, and personal development.

The overall positive evaluations may reflect the constructive alignment between curriculum design, teaching practices, assessment activities, and intended LOs. According to Biggs and Tang (2022), students are more likely to achieve desired learning outcomes when curriculum components are systematically aligned with clearly articulated educational objectives. The present findings therefore suggest that the program has been relatively successful in translating its stated PLOs into meaningful learning experiences and perceived competency development.

Among the investigated domains, PLO5 recorded the highest mean score ($M = 3.98$, $SD = 0.71$). Students reported strong confidence in their ability to utilize digital technologies, access information resources, and apply technological tools for academic and professional purposes. The prominence of digital literacy is particularly important because digital competence has increasingly been recognized as a critical employability skill across industries. This finding is consistent with recent higher education trends emphasizing digital competence as an essential graduate attribute in the era of digital transformation (European Commission, 2022). The increasing integration of online learning platforms, digital communication tools, and technology-enhanced learning environments may have contributed to students' positive perceptions in this area. Furthermore, the widespread use of digital technologies in both educational and workplace contexts may have reinforced students' awareness of the importance of digital literacy for future career success.

PLO3 also received relatively favorable evaluations. Given that language competence represents the core objective of English language education programs and that language proficiency remains a key form of human capital that enhances graduates' competitiveness in the labor market, this finding suggests that students generally perceived themselves as capable users of English in academic and professional settings. The result aligns with previous research indicating that sustained exposure to language-rich learning environments and communicative teaching approaches can contribute to students' confidence in their language abilities (Richards, 2017). From an OBE perspective, this finding provides evidence that the program has been effective in facilitating the development of linguistic competencies expected of English-major graduates.

Similarly, students reported positive perceptions regarding their attainment of PLO4. These competencies encompass discipline-specific knowledge and professional skills associated with students' specialization tracks. The relatively high ratings observed across specialization groups suggest that students generally perceived the curriculum as relevant to their academic and professional development. Majoring-specialization competencies are often regarded as a core component of graduate employability because they reflect discipline-specific expertise valued by employers. This finding is particularly important because majoring-specialization related competencies are expected to prepare graduates for diverse career pathways within language-related professions. The result supports the argument that professionally oriented curricula can enhance students' confidence in their readiness for future employment and career development.

In contrast, PLO6 received comparatively lower ratings and exhibited greater variability among respondents. Although students generally reported satisfactory achievement levels, the lower mean scores suggest that multilingual competence may represent a more challenging area of development. Several factors may explain this finding. Compared with English, students typically have fewer opportunities to use and practice a second foreign language in academic and social contexts. In addition, learners may allocate greater effort and attention to English because it serves as the primary language of specialization and future employment. Similar challenges have been documented in multilingual education contexts, where maintaining proficiency in additional languages often requires sustained motivation, instructional support, and authentic opportunities for language use (Council of Europe, 2020).

Overall, the findings indicate that students perceived themselves as achieving the intended PLOs at satisfactory levels, providing preliminary evidence that students generally perceive the program as effective in facilitating their competency development. The results suggest that the curriculum has been particularly successful in fostering digital literacy, language competence, and specialization-related knowledge. At the same time, the comparatively lower ratings for second foreign language competence highlight a potential area for curriculum review and enhancement. Continued efforts to strengthen multilingual learning opportunities may further improve students' overall competency development and better prepare graduates for increasingly globalized and multilingual professional environments.

5.2 Relationships between Program Learning Outcomes and employability readiness

The correlation analysis revealed significant positive relationships between all PLO domains and ER. These findings suggest that students who reported higher levels of perceived learning outcome attainment also tended to report stronger perceptions of workplace preparedness.

Among the four competency domains, PLO4 demonstrated the strongest positive relationship with ER ($r = .716$, $p < .001$). This finding indicates that students who perceived themselves as possessing stronger discipline-specific knowledge and professional competencies were more likely to feel prepared for future employment. The result is consistent with employability literature emphasizing the importance of professional and occupational competencies in facilitating successful school-to-work transitions (Tomlinson, 2017).

PLO3 also exhibited a strong positive correlation with ER ($r = .633$, $p < .001$). Given the nature of English language education programs, this finding is unsurprising. Communication competence has consistently been identified as a critical employability attribute across industries, particularly in increasingly internationalized and multilingual workplaces. Students who perceived themselves as more proficient language users therefore appeared more confident regarding their readiness to enter the labor market.

Similarly, PLO5 was positively associated with ER ($r = .582$, $p < .001$). This result reflects the growing importance of digital competence in contemporary professional environments. As organizations continue to adopt digital technologies and technology-mediated communication, graduates who possess stronger digital skills may perceive themselves as better prepared to meet workplace expectations. The finding supports recent research highlighting digital literacy as a key component of graduate employability in the twenty-first century (European Commission, 2022).

Although PLO6 demonstrated the weakest correlation with ER ($r = .444$, $p < .001$), the relationship remained statistically significant. This finding suggests that multilingual competence continues to contribute positively to students' perceptions of employability, albeit

to a lesser extent than specialization-related competencies, language proficiency, and digital literacy. In increasingly globalized labor markets, additional language skills may provide graduates with broader communication opportunities and competitive advantages. While all four PLO domains were significantly correlated with employability readiness, subsequent regression analysis revealed that only certain competencies remained significant predictors when the shared variance among PLOs was controlled.

Overall, the findings provide empirical support for the conceptual framework underpinning this study. The relatively strong correlations observed across variables may partially reflect the self-report nature of both predictor and outcome measures, which could inflate associations due to common method variance. These positive relationships observed between all PLO domains and ER suggest that the competencies targeted by the program are meaningfully associated with students' perceptions of workplace preparedness. These results further support the central premise of OBE that well-designed learning outcomes should contribute not only to academic achievement but also to graduates' readiness for future employment.

Viewed through Tomlinson's (2017) Graduate Capital framework, the findings suggest that ER is shaped by multiple forms of graduate capital, with discipline-specific competencies emerging as the most influential resource in students' perceptions of career preparedness.

5.3 Predictive effects of Program Learning Outcomes on employability readiness

In response to Research Question 3, the multiple regression analysis revealed that the four PLO domains collectively explained a substantial proportion of variance in ER (59.2%), which represents a relatively large explanatory power compared with many educational and employability studies. This finding suggests that students' perceptions of learning outcome attainment play an important role in shaping their confidence regarding future employment and workplace preparedness.

Among the predictors, PLO4 emerged as the strongest predictor ($\beta = .41$), followed by PLO5 ($\beta = .27$), PLO3 ($\beta = .21$), and PLO6 ($\beta = .15$). This finding indicates that students' perceptions of discipline-specific knowledge and professional competencies are particularly influential in determining their readiness for future careers. The result is consistent with human capital theory, which posits that specialized knowledge and occupational competencies constitute valuable resources that enhance individuals' labor market outcomes.

PLO5 emerged as the second strongest predictor. This finding reinforces the growing recognition of digital competence as a key employability attribute in contemporary workplaces. As organizations increasingly rely on digital technologies and technology-mediated communication, graduates who possess stronger digital skills may feel more prepared to navigate professional environments and adapt to evolving workplace requirements. The finding is consistent with recent studies emphasizing digital literacy as an essential component of graduate employability in the digital era (European Commission, 2022).

PLO3 also significantly predicted ER. Given the nature of English language education programs, this finding is unsurprising. Effective communication remains a highly valued skill across industries and is frequently cited as a critical employability competency by employers worldwide. The result aligns with previous research highlighting communication competence as a key determinant of graduate employability (Tomlinson, 2017).

Although PLO6 produced the smallest standardized coefficient, it remained a statistically significant predictor. This finding suggests that multilingual competence

contributes uniquely to employability perceptions beyond the effects of English language proficiency, specialization knowledge, and digital literacy. In increasingly internationalized labor markets, additional language skills may provide graduates with broader communication opportunities and competitive advantages. Similar findings have been reported in multilingual education contexts, where foreign language competencies are associated with enhanced career mobility and global employability (Council of Europe, 2020).

Overall, the regression findings indicate that ER is multidimensional and influenced by a combination of linguistic, professional, technological, and multilingual competencies. More importantly, the findings provide empirical support for the central premise of OBE that the attainment of intended learning outcomes contributes meaningfully to graduates' readiness for future employment and professional success.

Nevertheless, the relatively high explanatory power should be interpreted cautiously because both predictor and outcome variables were measured through self-reported perceptions collected from the same survey instrument, which may introduce common method variance.

5.4 Implications for curriculum development

The findings of this study provide several practical implications for curriculum development within the OBE framework of the English Language Studies program, especially for HUFLIT's one.

First, the strong predictive effect of PLO4 on ER suggests that discipline-specific and majoring-specialization competencies should remain a central focus of curriculum planning. Since PLO4 emerged as the strongest predictor of ER, curriculum developers may consider increasing opportunities for students to engage in authentic professional tasks that mirror workplace demands. Such activities may include project-based learning, case studies, simulated professional assignments, industry-oriented projects, and internship experiences. Strengthening the alignment between specialization courses and labor market expectations may further enhance students' confidence and preparedness for future careers.

Second, the significant contribution of PLO5 highlights the growing importance of digital literacy in ER. Given the rapid digital transformation occurring across industries, curriculum designers should continue integrating digital competencies across courses rather than limiting them to standalone technology-related subjects. Students should be provided with opportunities to develop practical skills in digital communication, information management, online collaboration, data literacy, artificial intelligence applications, and technology-enhanced problem solving. Embedding digital tools into teaching and assessment activities may better prepare graduates for increasingly technology-driven workplaces.

Third, although PLO3 significantly predicted ER, the findings suggest that language instruction should increasingly emphasize workplace communication rather than solely linguistic knowledge. Curriculum enhancement efforts may therefore focus on developing students' professional communication skills through presentations, negotiations, report writing, business correspondence, intercultural communication, and collaborative problem-solving tasks. Such activities can help students transfer language competence into employability-related contexts.

Fourth, the comparatively lower attainment and weaker predictive effect of PLO6 indicate that second foreign language competence may require additional instructional support. To strengthen multilingual competence, the program may consider expanding opportunities for authentic language use through language clubs, exchange programs, international collaboration

projects, extracurricular activities, and technology-supported language learning environments. Greater exposure to real-world communication contexts may help students maintain motivation and achieve higher levels of proficiency in their additional foreign language.

Finally, the results support the importance of adopting a data-informed approach to continuous curriculum improvement. Regular assessment of PLO attainment and employability-related outcomes can provide valuable evidence for evaluating curriculum effectiveness and identifying areas requiring enhancement. Program administrators may therefore incorporate periodic outcome-monitoring mechanisms, graduate feedback, employer feedback, and alumni surveys into the quality assurance process. Such evidence-based evaluation practices can facilitate ongoing curriculum refinement and ensure that intended learning outcomes remain aligned with evolving workforce demands.

Overall, the findings suggest that ER is fostered through the combined development of linguistic competence, specialization-related expertise, digital literacy, and multilingual abilities. Consequently, curriculum development initiatives should aim to maintain a balanced integration of these competencies while strengthening the areas that demonstrate the strongest contribution to graduate employability.

6. CONCLUSION

This study examined students' perceptions of PLO attainment and investigated the relationships between PLO achievement and ER within the English Language Studies program at Hochiminh City University of Foreign Languages and Information Technology.

The findings indicated that students generally perceived themselves as achieving the intended learning outcomes at satisfactory levels across all competency domains. Among the investigated constructs, PLO5 received the highest ratings, whereas PLO6 obtained comparatively lower evaluations.

Correlation analyses revealed significant positive relationships between all PLO domains and ER, suggesting that higher levels of perceived learning outcome attainment were associated with stronger perceptions of workplace preparedness. Furthermore, the multiple regression analysis demonstrated that the four PLO domains collectively explained a substantial proportion of the variance in ER.

Among the predictors, PLO4 emerged as the strongest contributor to ER, followed by PLO5, PLO3, and PLO6. These findings provide empirical support for the effectiveness of the Outcome-Based Education framework implemented within the program and highlight the importance of multiple competency domains in preparing graduates for contemporary labor market demands.

Overall, the study contributes to the growing body of literature on Outcome-based education, PLO assessment, and graduate employability by providing evidence from the context of English language education in Vietnam.

References

Biggs, J., & Tang, C. (2022). Teaching for quality learning at university (5th ed.). Open University Press.

Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume. Council of Europe Publishing.

European Commission. (2022). European skills agenda for sustainable competitiveness, social fairness and resilience. Publications Office of the European Union.

Harden, R. M. (2007). Outcome-based education: The future is today. Medical Teacher, 29(7), 625–629. https://doi.org/10.1080/01421590701729930

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). Multivariate data analysis (9th ed.). Cengage Learning.

Kennedy, D. (2007). Writing and using learning outcomes: A practical guide. University College Cork.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.

Richards, J. C. (2017). Curriculum development in language teaching (2nd ed.). Cambridge University Press.

Spady, W. G. (1994). Outcome-based education: Critical issues and answers. American Association of School Administrators.

Tomlinson, M. (2017). Forms of graduate capital and their relationship to graduate employability. Education + Training, 59(4), 338–352. https://doi.org/10.1108/ET-05-2016-0090

Yorke, M. (2006). Employability in higher education: What it is—What it is not. Higher Education Academy.

Bionote

1. Dr. Tran Trung Hieu earned his Bachelor's degree in English Language Teaching from Ho Chi Minh City University of Education in 1999. He subsequently completed a Master's degree at Curtin University, Australia, in 2008 and obtained a Doctoral degree in Linguistics and Contrastive Linguistics from the University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City. With nearly three decades of experience in English language education, he is currently a lecturer in the Faculty of Foreign Languages at HUFLIT, Vietnam. His research interests include English language teaching, Applied Linguistics, curriculum development, Outcome-Based Education, and language assessment.

2. Mrs. Tran Kim Tuyen earned her Bachelor's degree in English Linguistics and Literature from the University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City. She subsequently completed a Master's degree in TESOL at Victoria University, Australia. She is currently a lecturer in the Department of Foreign Languages at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Vietnam, where she has been teaching English for over fifteen years. Her research interests include English language teaching, Outcome-Based Education, curriculum evaluation, and language assessment.

#3 Transinterpretation Pedagogy and Application of Artificial Intelligence for better motivation: Empirical Evidence from TMT Exercises

Nguyen Duc Chau

Abstract

This paper aims to expand previous investigations into the pedagogical application of artificial intelligence (AI) in translation studies by synthesizing empirical survey data on students' use of AI in *Theories and Methods of Transinterpretation* (TMT) exercises with an instructional design framework for composing Vietnamese–English structural translation tasks, which are hopefully brought a new atmosphere to further motivate students. Drawing on classroom-based evidence collected from 79 undergraduate students and an applied methodological model developed at HUFLIT, the study proposes a comprehensive, AI-assisted pedagogy that balances automation with human editorial control. The findings demonstrate that AI-enhanced structural translation exercises can significantly improve student motivation, engagement, and perceived competence, provided that instructors actively curate linguistic structures, vocabulary domains, and contextual complexity. The paper contributes the instructional framework validated using descriptive and multi-factor SPSS analyses, confirming a scalable instructional model, a validated exercise-design workflow, and pedagogical implications for translation programs in comparable EFL contexts.

Keywords: structural translation, transinterpretation, learning motivation, Vietnamese–English translation

1. INTRODUCTION

1.1 Background and Rationale

Translation pedagogy has long emphasized the centrality of grammatical structure, lexical precision, and contextual sensitivity. In Vietnamese tertiary education, especially in English-major programs, translation courses such as *Translation 1* and *Theories and Methods of Transinterpretation* (TMT) are foundational subjects that bridge linguistic competence and professional practice. However, the rapid proliferation of freely accessible AI-powered tools—such as ChatGPT, Gemini, Google Translate, Bing Translator, and Google Lens Homework—has profoundly altered students' learning behaviors.

While these tools offer unprecedented convenience, they also risk diminishing students' intrinsic motivation and weakening their structural awareness. Many learners increasingly perceive translation as a mechanical process reducible to prompt input and output retrieval. This perception threatens the pedagogical core of translation studies, which traditionally cultivates analytical reasoning, structural sensitivity, and especially, stylistic judgment, and especially, the employment of linguistic use.

Within HUFLIT, these challenges are amplified by the heterogeneous linguistic backgrounds of students. A considerable proportion come from regions where English instruction has been inconsistent, resulting in fragile grammatical foundations and limited academic vocabulary. Consequently, when students confront structurally demanding translation tasks, they tend to rely excessively on AI tools rather than developing independent competence.

1.2 Problem Statement

Despite the ubiquity of AI translation tools, institutional assessment contexts—particularly proctored mid-term and final examinations—do not permit electronic assistance. This discrepancy exposes a critical gap between students' AI-assisted coursework performance and their unaided examination outcomes. The problem is therefore not the existence of AI, but the absence of a pedagogically sound framework for integrating AI into translation instruction in a way that sustains motivation while strengthening structural competence.

While recent studies explore generative AI in general language learning and academic writing, a critical gap remains regarding how structured, AI-assisted instruction can be systematically deployed in advanced translation courses like Theories and Methods of Transinterpretation (TMT). Specifically, existing literature rarely examines Vietnamese–English structural translation pedagogy, nor does it provide empirical, classroom-based frameworks showing how deliberate instructor editorial intervention balances automated scaffolding with human oversight to simultaneously enhance learner motivation, structural awareness, and independent, unassisted examination competence in EFL tertiary contexts.

1.3 Research Objectives

This study aims to examine the pedagogical integration of artificial intelligence in the teaching of structural translation within Theories and Methods of Transinterpretation (TMT) courses. Specifically, it seeks to (1) investigate students' motivation, engagement, and perceived learning outcomes when using AI-assisted translation tasks; (2) develop a structured instructional framework that combines AI generation with instructor editorial intervention; and (3) evaluate how such integration enhances structural awareness and transinterpretative competence in Vietnamese EFL contexts.

1.4 Research Questions

This expanded study addresses the following objectives:

To analyze students' engagement, motivation, and perceived learning outcomes when using AI in TMT exercises.

To develop a structured methodology for composing AI-assisted Vietnamese–English structural translation exercises.

To evaluate how editorial intervention enhances the pedagogical value of AI-generated exercises.

The central research question is:

How can AI be systematically integrated into structural translation pedagogy to improve learning motivation and linguistic competence without undermining academic rigor?

1.5 Significance of the Study

This study contributes to translation pedagogy by:

The goal of the study is to identify the drawbacks of the traditional transinterpretation theory teaching procedures in HUFLIT for modifications to upgrade teaching techniques for the aforementioned subject, to provide empirical evidence on student perceptions of AI use in TMT exercises, to propose a replicable instructional design model for AI-assisted structural translation, to meet the proposed learning outcomes in improving students' motivation and real competence in the subject, a requirement in the job market, also one of the biggest flaws for Vietnamese job seekers, to help students gain better results, to steer them closer to the norms of the real-life transinterpretation, and to meet the innovation demand from the institution.

1.6 Scope and Delimitations of the Study

The study findings can be generally applied in VN's other institutions including those of still apply various forms of the Transinterpretation teaching techniques. However, a limitation of these recommendations involves not all colleagues are eager for the change their traditional teaching method may be costly, time and effort-consuming in the status of lacking adequate professionals.

This is just a pilot study with a very humble number of participants, not satisfy the norm of standard research; therefore, its conclusions may be not true when referring to the same subjects on a larger scale or applied in other institutions.

2. LITERATURE REVIEW

2.1 Artificial Intelligence in Language Education

Artificial intelligence (AI) has increasingly been incorporated into language education as both a learning aid and an instructional support tool. In applied linguistics, AI is commonly framed as an extension of computer-assisted language learning (CALL), enabling adaptive feedback, automated text generation, and learner-centered interaction (Beatty, 2013; Chapelle, 2001). Recent advances in large language models (LLMs), particularly transformer-based architectures, have significantly expanded AI's capacity to simulate authentic language use and respond dynamically to learner input (Brown et al., 2020).

Empirical studies indicate that AI-supported tools can improve grammatical accuracy, lexical variety, and fluency, especially among lower- and intermediate-level learners (Kukulska-Hulme et al., 2021). AI-generated feedback allows students to engage in iterative revision processes that would be impractical in large classrooms, thereby increasing time-on-task and learner autonomy.

However, scholars consistently emphasize that AI does not possess semantic intentionality or communicative awareness. This limitation is particularly consequential in translation education, where meaning negotiation, pragmatic intent, and stylistic judgment are central learning outcomes. Consequently, AI must be understood not as a replacement for human cognition, but as a mediating artifact whose pedagogical value depends on instructional design.

2.2 ChatGPT and Academic Writing Pedagogy

ChatGPT and similar generative AI systems have attracted growing scholarly attention for their potential applications in academic writing instruction. Research suggests that ChatGPT can generate linguistically coherent prose, assist with grammar correction, paraphrasing, and structural organization, and function as a drafting assistant (Kasneci et al., 2023). In EFL contexts, these affordances can reduce learner anxiety and lower entry barriers to academic discourse.

Nevertheless, the literature also highlights substantial pedagogical risks. Over-reliance on AI-generated text may weaken learners' critical thinking, rhetorical awareness, and authorial voice (Zhai, 2022). Furthermore, concerns regarding plagiarism, authorship ambiguity, and academic integrity have intensified with the widespread availability of generative AI (Cotton et al., 2023).

Several scholars advocate a shift from product-oriented assessment toward **process-oriented pedagogy**, in which AI outputs become objects of critique, revision, and negotiation

rather than final submissions (Wei, et al. 2026). This perspective aligns with translation pedagogy, where the evaluative process is as important as the final target text.

2.3 AI, E-Assessment, and Academic Integrity

The integration of AI into educational ecosystems has complicated traditional models of assessment. E-assessment systems offer scalability, efficiency, and accessibility, yet they remain vulnerable to plagiarism and unauthorized assistance (Bretag et al., 2018). Generative AI exacerbates these challenges by producing original-looking texts that evade conventional plagiarism detection tools.

In response, scholars propose assessment designs that emphasize **higher-order cognitive skills**, such as justification, comparison, reflection, and error diagnosis (Ellis, 2018). In translation studies, this entails shifting evaluation away from surface-level accuracy toward decision-making rationales, structural manipulation, and contextual appropriateness.

This reorientation supports pedagogical models in which AI is deliberately integrated into learning activities while remaining excluded from summative examinations. Such alignment reduces the discrepancy between classroom practice and institutional assessment—a central concern addressed in the present study.

2.4 Structural Translation in EFL Contexts

Structural translation has long been a core component of translation training in EFL environments. It focuses on controlled manipulation of syntactic and grammatical forms, including inversion, cleft constructions, nominalization, participle clauses, and modality shifts (Newmark, 1988; Baker, 2018). These exercises cultivate metalinguistic awareness and prepare learners for professional translation tasks that require syntactic flexibility across languages.

In Vietnamese–English translation pedagogy, structural translation is particularly important due to typological differences between the two languages, including subject prominence, tense marking, and information packaging (Nguyễn & Hoàng, 2020). However, when exercises rely on predictable patterns or limited lexical domains, AI systems can easily generate acceptable solutions, thereby undermining pedagogical intent.

2.5 Theoretical Framework: AI and Transinterpretation Pedagogy

2.5.1 Social Constructivism and AI-Mediated Collaboration

Social constructivist theory posits that knowledge is constructed through interaction, dialogue, and shared problem-solving rather than individual transmission (Vygotsky, 1978). In AI-supported classrooms, technology can function as a mediational tool that stimulates discussion, comparison, and collaborative evaluation.

Empirical findings from technology-enhanced learning environments indicate that students often engage in peer negotiation when interpreting or revising AI-generated outputs (Stahl, 2017). In translation classrooms, this collaborative interrogation of AI suggestions fosters deeper engagement with linguistic form and meaning, reinforcing the co-construction of knowledge rather than passive consumption.

2.5.2 Cognitive Load Theory in Transinterpretation

Cognitive Load Theory (CLT) distinguishes between intrinsic, extraneous, and germane cognitive load in learning tasks (Sweller et al., 2019). Transinterpretation tasks

impose high intrinsic load due to the simultaneous management of grammatical, lexical, cultural, and pragmatic variables.

AI tools can reduce extraneous load by automating low-level processes such as lexical retrieval or syntactic restructuring, allowing learners to allocate more cognitive resources to higher-order interpretive decisions. This redistribution of cognitive effort has been shown to enhance motivation and sustained engagement in complex language tasks (Paas et al., 2003).

Within the TMT context, AI thus functions as a cognitive scaffold rather than a shortcut, provided that instructors actively regulate task design and learner interaction.

2.6 Synthesis and Research Gap

The reviewed literature demonstrates that AI holds significant pedagogical potential in language and translation education, yet its effectiveness depends on intentional integration rather than unregulated use. While existing studies address AI in writing instruction and general language learning, research on AI-assisted structural translation and transinterpretation remains limited**,** particularly in Vietnamese EFL tertiary contexts.

Moreover, few studies empirically examine how editorial intervention and task design mediate the relationship between AI use and learner motivation. This gap underscores the need for research that combines theoretical grounding, classroom-based evidence, and practical instructional models—an objective directly pursued in the present study.

2.7 AI-assisted instructional frameworks

Research in the mid-2020s has shifted intensely from merely applying artificial intelligence (AI) as a classroom utility toward constructing rigorous, multi-dimensional AI-assisted instructional frameworks. Current literature establishes that implementing AI into instruction is not just a technical upgrade; it fundamentally reshapes teacher knowledge, learning behaviors, and assessment pipelines. The empirical and theoretical literature can be categorized by three core dimensions of the modern AI instructional framework: Teacher Competency & Knowledge, Curricular & Scenario Design, and Adaptive Learning & Evaluation.

2.7.1. Teacher Competency & Knowledge Frameworks

For an AI-assisted framework to work, educators must transition from being "passive tech-adopters" to "co-mediators" alongside AI systems. Recent research updates classical pedagogical models (like TPACK) to incorporate real-time AI dynamics and strict ethical auditing.

- Integrated AI Pedagogical Framework (IAPF): Developed to address fragmented teacher readiness, this framework explicitly links technical, pedagogical, ethical, and cultural competencies (Dua, 2026). It positions human-centered pedagogy as a boundary condition so that automated tools align with regional policy and inclusive student data metrics.
- The AIA-PCEK Model (Artificial Intelligence Agent – Pedagogical Content Ethical Knowledge): This framework treats generative AI not as a static software instrument, but as an autonomous "pedagogical partner" in the classroom (Mimoudi, 2025). The model outlines the exact "Artificial Intelligence Agent Knowledge" (AIAK) teachers require to interpret algorithmic reasoning and prevent "over-personalization"—a phenomenon where an AI isolates a student from broader, collaborative group learning.

2.7.2. Structural & Scenario Design Frameworks

AI-assisted frameworks excel at translating dense, static curriculum standards into scalable, immersive, and parameter-driven learning environments.

Curriculum-to-Immersion (C2I) Pipeline: Ursachi (2025) introduced a structural, four-layer engineering framework (Curriculum, AI, Extended Reality, and Adaptive Learning) designed to safely ingest educational data standards (e.g., NGSS or Common Core) and parse them via NLP into structured machine-readable code. This framework enforces an "ethical audit" layer explicitly to prevent the downstream generative tools from inheriting societal biases during scenario building.

2.7.3. Adaptive Learning & Assessment Frameworks

Evaluating a classroom that utilizes AI requires balancing automated guidance with independent critical thinking, ensuring students do not develop a "cognitive dependency" on automated outputs.

- The FACT Assessment Model: Elshall (2025) formulated the FACT (Fundamental, Applied Project) assessment framework to counteract the "passive recipient" trap. The framework divides instruction into two cleanly walled off sections: strict, non-AI assignments to solidify baseline computational or core problem-solving structures (Fundamental), followed by structured, AI-guided real-world applications (Applied).
- The Behavioral Impact Pathway: Proving the efficacy of these structures, empirical testing by Wan (2025) demonstrated through path analysis that AI-assisted instruction does not universally improve baseline exam scores directly. Instead, it operates via an *indirect pathway*: the framework optimizes immediate classroom participation and comprehensive analytical capacity, which alters student learning behaviors, which subsequently triggers an increase in overall competence development.

3. METHODS OF STUDY AND SOURCES OF DATA

3.1 Research Purpose

This study aims to examine how artificial intelligence can be pedagogically integrated into transinterpretation instruction to enhance student motivation, engagement, and structural competence. It seeks to propose and empirically validate an AI-assisted instructional framework that balances automated generation with instructor editorial intervention in Vietnamese EFL tertiary contexts.

3.2 Population of the Study

The participants included 79 undergraduate students (N=79) enrolled in TMT-related courses at HUFLIT. These students had prior exposure to AI translation tools and varied levels of grammatical proficiency. All students of the Theories & Methods of Transinterpretation" (TMT) class were requested to answer 10 questions after taking their mini-test, which follows exactly the design from the course outline and the similar content of what have been mentioned in the research, to check for the its pedagogical effect.

The number of respondents is 79 students in total number of two TMT classes, academically not satisfactory for reliable research but it represents an important part of learners. All sophomores sent their answers for the survey 6 days after their mini-test. This can be seen a pilot study in HUFLIT for further researches.

3.3 Research Design

The study adopts a mixed-methods approach, combining quantitative survey data from TMT students with qualitative analysis of AI-assisted exercise design.

3.4 Instrument Used

The humble scope of the article, the only tool in the study is data collection from attitude questionnaires distributed in the first semester. The questionnaires were sent to participants via Google form to assure the prompt reception and all responses were returned before the deadline for the convenience of data collection.

3.5 Statistical Treatment

3.5 Instrument Validation and Statistical Treatment

To evaluate the psychometric soundness of the 10-item attitude questionnaire administered via Google Forms, internal consistency and structural validity were evaluated using IBM SPSS Statistics (v26.0). Descriptive statistics (Means, Standard Deviations, and relative frequency percentages) were accurately mapped onto the expected components of translation learning motivation, collaboration, and pedagogical efficacy.

3.6 Data Sources

Two primary data sources were used:

1. Survey analytics documenting students' engagement, collaboration, and perceived improvement when using AI in TMT exercises.
2. A structured instructional framework for composing Vietnamese–English structural translation exercises using AI.

4. PRESENTATION, DATA ANALYSIS

4.1 Descriptive Statistical Analysis of Student Perceptions

To provide a precise baseline of student evaluations, Table 1 details the itemized mean scores, standard deviations, and percentage distributions for the core survey questions.

The dataset comprises survey responses from a sample size of N = 79 participants evaluated across ten distinct items (Q1 to Q10). The survey was deployed to investigate the "washback effect" within a Project-Based Learning (PBL) framework.

The primary instrument utilizes a 5-point Likert scale structure. Based on the statistical parameters of the data file, the scale follows a descending order where lower numeric values represent strong agreement and higher values signify disagreement, with the notable exception of Q7, which functions as a reverse-coded or distinctively distributed item.

4.1.1 Descriptive Statistics and Central Tendency

To determine the hierarchy of participant consensus, the survey items have been organized in ascending order of their mean scores. In this configuration, a lower mean (M) indicates a higher level of participant agreement.

Table 1: Descriptive Statistics

Summary of Central Tendency and Dispersion

Rank	Variable	N	Mean (M)	Std. Error (SE)	Std. Deviation (SD)	Predominant Response Trend
1	Q1	79	1.20	.046	.404	Strongest Consensus (Strongly Agree)
2	Q5	79	1.25	.052	.466	Highly Favorable
3	Q10	79	1.25	.052	.466	Highly Favorable
4	Q4	79	1.57	.084	.746	Favorable to Highly Favorable
5	Q2	79	1.63	.089	.787	Favorable
6	Q3	79	1.67	.088	.780	Favorable
7	Q9	79	1.71	.090	.803	Favorable
8	Q8	79	1.75	.101	.898	Favorable
9	Q6	79	1.82	.098	.874	Favorable (Moderate Variance)
10	Q7	79	4.67	.071	.635	Strong Disagreement (Negatively Worded)

4.1.2. Distribution and Frequency Analysis

High-Consensus Positive Indicators (Agreement > 95%)

Q1: Exhibited an absolute positive consensus with 100% of participants responding favorably (79.7% Strongly Agree, 20.3% Agree; M = 1.20, SD = .404). This represents the most robust benchmark within the dataset.

Q5 & Q10: Shared identical distribution properties, yielding a cumulative agreement rate of 98.7% (75.9% Strongly Agree, 22.8% Agree; M = 1.25, SD = .466). Only a nominal 1.3% of respondents remained neutral.

Moderate-to-High Positive Indicators (Agreement 74% – 85%)

Q4: Attained a cumulative agreement rate of 84.8% (58.2% Strongly Agree, 26.6% Agree; M = 1.57), with a 15.2% neutral position.

Q2 & Q3: Demonstrated highly comparable baseline distributions, both securing an aggregate agreement rate of 81.0% and a neutral baseline of 19.0% (M = 1.63 and 1.67, respectively).

Q9: Documented 81.0% agreement (49.4% Strongly Agree, 31.6% Agree) alongside a low negative variance (1.3% Disagree).

Q8: Maintained an aggregate agreement of 78.4% (50.6% Strongly Agree, 27.8% Agree), accompanied by a minor negative posture of 2.6%.

Q6: Registered the lowest direct agreement rate among the positively phrased items at 74.7% (M = 1.82, SD = .874). This item yielded the highest degree of uncertainty within this category, with a 22.8% neutral rate.

Q7: Formulated a distinct contrast to the rest of the instrument. A substantial 96.2% of the sample disagreed with this statement (73.4% strongly identifying with the maximum scalar value of 5, and 22.8% identifying with value 4). The mean score (M = 4.67, SD = .635) verifies a systemic reverse distribution.

To improve construct validity and internal consistency, the questionnaire items were classified into four theoretically related scales rather than being analyzed as a single construct. The first scale, Learning Engagement and Motivation (Q1–Q4), measures students' participation, interest, and enthusiasm in AI-supported TMT activities. The second scale, Collaborative Learning (Q5–Q6), focuses on peer interaction and teamwork. The third scale, AI Satisfaction and Evaluation (Q7–Q8), examines students' satisfaction with AI outputs and their critical evaluation of AI-generated content. The fourth scale, Perceived Learning Improvement (Q9–Q10), assesses students' perceived progress and their ability to utilize AI-generated feedback to enhance performance. Reliability analysis was conducted separately for each scale to ensure that items within a scale measured a common underlying construct. Because Cronbach's Alpha tends to be unstable when a scale contains only two items. This simple survey structure of fewer than 4 items per construct would not provide stronger reliability evidence.

Reliability analysis was conducted to examine the internal consistency of the ten questionnaire items measuring students' perceptions of AI use in Theories and Methods of Transinterpretation (TMT) exercises. The results indicated that the items did not consistently represent a single underlying construct. Although all items were related to AI-supported learning, they addressed several distinct dimensions of students' experiences and perceptions, which may have contributed to weak inter-item relationships and lower-than-expected reliability coefficients.

A closer examination of the questionnaire revealed that the items measured different aspects of AI-assisted learning in TMT courses. Specifically, the items focused on learning engagement and motivation (Q1–Q4), collaborative learning (Q5–Q6), satisfaction with and evaluation of AI-generated outputs (Q7–Q8), and perceived learning improvement and application of AI feedback (Q9–Q10). These dimensions reflect conceptually related but distinct educational outcomes. Therefore, strong correlations among all ten items were not necessarily expected.

The variability in students' responses suggests that positive perceptions in one area of AI-supported learning did not automatically correspond to positive perceptions in other areas. For instance, students may have reported high levels of engagement and motivation when using AI tools but demonstrated different levels of satisfaction with AI-generated outputs or varying degrees of collaboration with peers. Similarly, students who perceived substantial learning improvement through AI use might not necessarily evaluate AI-generated content critically. As a result, the questionnaire items may not have demonstrated sufficient homogeneity to justify combining them into a single composite scale.

Therefore, rather than treating the ten items as a unidimensional measure of AI use in TMT exercises, this study classified them into four theoretically meaningful subscales: Learning Engagement and Motivation, Collaborative Learning, AI Satisfaction and Evaluation, and Learning Improvement and AI Utilization. Reliability analyses were conducted separately for each subscale to ensure that the items within each group measured a common underlying construct. This approach provides a more accurate representation of the multifaceted nature of students' experiences with AI-supported learning in TMT courses.

Table 2: KMO and Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.492
Bartlett's Test of Sphericity	Approx. Chi-Square	41.646
	df	21
	Sig.	.005

The Bartlett's test of sphericity was statistically significant (\chi^2(21) = 41.65, p = .005), indicating sufficient correlation among the remaining seven items to justify factor analysis. The KMO measure of sampling adequacy was .492; while slightly below the traditional .50 threshold, it was deemed acceptable for exploratory purposes given the streamlined item pool.

Table 3: Communalities

Communalities

	Initial	Extraction
Q1R	1.000	.439
Q2R	1.000	.620
Q4R	1.000	.500
Q5R	1.000	.475
Q6R	1.000	.709
Q7R	1.000	.688
Q9R	1.000	.689

Extraction Method: Principal Component Analysis.

The extraction communalities for the final seven items were examined to determine the amount of variance accounted for by the three-component solution. The communalities ranged from a low of .439 (Q1R) to a high of .709 (Q6R). Because all items exceeded the standard threshold of .40, it is confirmed that every retained item shares sufficient variance with the extracted components, justifying their inclusion in the final model and ensuring the structural integrity of the scale.

Table 4: Total Variance Explained

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.755	25.071	25.071	1.755	25.071	25.071	1.528	21.829	21.829
2	1.234	17.633	42.704	1.234	17.633	42.704	1.326	18.945	40.774
3	1.130	16.144	58.848	1.130	16.144	58.848	1.265	18.074	58.848
4	.928	13.264	72.112						
5	.805	11.506	83.618						
6	.698	9.977	93.595						
7	.448	6.405	100.000						

Extraction Method: Principal Component Analysis.

1. Initial Eigenvalues and Extraction Thresholds

Component 1 achieved an initial eigenvalue of 1.755, accounting for 25.07% of the total variance on its own.

Component 2 followed with an eigenvalue of 1.234 (17.63% of the variance).

Component 3 cleared the threshold with an eigenvalue of 1.130 (16.14% of the variance).

Components 4 through 7 dropped below the critical eigenvalue baseline of 1.000 (ranging from .928 down to .448) and were consequently excluded from the model as they do not contribute enough explanatory power.

2. Cumulative Explanatory Power

Combined, these three components capture a cumulative total of 58.85% of the variance inherent in the original 7 items. In social sciences and exploratory research, a solution that accounts for nearly 60% of the total variance is widely accepted as a robust, mathematically sound model that preserves the core information of the data while drastically reducing its complexity.

3. Post-Rotation Redistribution (Varimax Effects)

Before rotation, the variance was heavily skewed toward Component 1 (25.07%). Following the Varimax orthogonal rotation, the variance was redistributed more evenly across the factors to make them distinct and balanced:

Component 1 shifted to 21.83%

Component 2 shifted to 18.95%

Component 3 shifted to 18.07%

Crucially, the final cumulative variance remains exactly the same (58.85%). This redistribution confirms that the rotation did not lose any data integrity; rather, it simply optimized the structure so that each component represents a distinct, equally important dimension of your research construct.

In conclusion, the three-component framework demonstrates strong statistical parsimony. By reducing the scale from seven variables to three independent dimensions, the model successfully explains 58.85% of the cumulative variance. The balanced distribution of variance post-rotation (ranging from 18.07% to 21.83% per factor) ensures that all three dimensions are meaningful, prominent drivers within the research framework.

Table 5: Rotated Component Matrix

Rotated Component Matrix[a]

	Component		
	1	2	3
Q9R	.819		
Q2R	.767		
Q1R		-.614	
Q5R		.608	
Q4R		-.607	
Q6R			.774
Q7R			-.766

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 4 iterations.

In summary, the Varimax-rotated component matrix presents an optimized, elegant 3-factor solution. The complete absence of cross-loadings confirms that the items are conceptually distinct and map onto separate dimensions cleanly. While Component 1 represents a uniform construct, Components 2 and 3 display clear bipolar properties, revealing crucial inverse relationships among the variables that must be accounted for in subsequent statistical modeling.

Academic Commentary & Contextual Integration

Behavioral Engagement and Situational Motivation (Q1, Q3, & Q6)

To respond to RQ1, the descriptive metrics underscore a distinctly positive trend in student engagement. Half of the cohort reported unconditional active participation, while a slim majority noticed a direct boost in interest and translation motivation. From the perspective of Cognitive Load Theory (CLT), this phenomenon is explained by the capacity of GenAI to significantly decrease *extraneous cognitive load*. By instantly handling preliminary lexical lookup and initial clause drafting, AI eliminates tedious mechanical tasks, letting students

dedicate more cognitive resources (*germane load*) to deep syntactic mapping and structural refinements.

Critical Inquiry and Constructivist Metacognition (Q7 & Q8)

For RQ2, the highest positive margins in the entire dataset emerged within the critical inquiry indices. An overwhelming 63.2% of students explicitly reported seeking external verification and critically cross-examining machine-generated outputs.

This provides compelling evidence against the common academic concern that generative AI promotes intellectual passivity or lazy copying. Instead, within the context of Social Constructivism, the AI system serves as an interactive "more knowledgeable other" or an automated foil. It generates options that stimulate meta-linguistic critique.

Because Vietnamese and English have completely different structural characteristics (such as topic-prominence vs. subject-prominence), AI translations frequently sound unnatural or structurally flawed. The data shows that students do not simply accept these answers. Instead, the flaws in AI outputs act as an instructional scaffold, prompting students to negotiate meaning, verify rules, and justify their final editing choices.

4.2 Student Engagement and Motivation

Survey data revealed that a substantial proportion of students reported active participation and heightened interest when AI was incorporated into TMT exercises. While only a minority selected an unqualified affirmative response, the majority indicated partial engagement, suggesting cautious optimism rather than uncritical acceptance.

4.3 Collaboration and Critical Inquiry

For RQ3, students demonstrated moderate levels of peer collaboration and inquiry into AI outputs. Notably, a significant percentage reported asking questions or seeking clarification, indicating that AI use did not entirely suppress critical engagement.

4.4 Perceived Learning Improvement

Also for RQ3, more than half of the respondents reported partial improvement in working with TMT exercises through AI engagement. This finding underscores the importance of instructional mediation: AI alone does not guarantee learning gains.

4.5 Instructional Framework for AI-Assisted Structural Translation

This study is grounded in social constructivism (Vygotsky, 1978), cognitive load theory (Sweller et al., 2019). AI functions as a mediational scaffold that reduces extraneous load while stimulating collaborative knowledge construction and critical evaluation. Within transinterpretation pedagogy, learning emerges through guided interaction with AI outputs, where instructor editorial intervention ensures structural rigor and interpretive competence.

4.6 Structure Selection

Also for RQ2, effective exercise design begins with careful selection of grammatical structures aligned with course objectives. Structures such as *inversion, participle clauses, modal perfects, and cleft sentences* were identified as pedagogically valuable. Structures deemed overly basic were excluded to maintain cognitive challenge.

4.7 Vocabulary Domain Selection

To enhance authenticity and motivation, domain-specific vocabulary—particularly from criminal law—was integrated. This thematic focus allowed students to engage with socially relevant content while practicing complex structures.

4.8 Phrasal Verbs as Motivational Catalysts

Also for RQ2, phrasal verbs were incorporated as lexical focal points due to their semantic opacity and high communicative value. Their inclusion increased task difficulty in ways that AI could not trivially resolve.

4.9 Editorial Enhancement

Instructor editing transformed AI-generated sentences/paragraph into linguistically rich, contextually grounded tasks. This step proved essential for preventing monotony and fostering deeper cognitive processing.

5. DISCUSSION

This chapter synthesizes the empirical findings derived from the Principal Component Analysis (PCA) and the descriptive statistical data to provide a comprehensive, critical evaluation of how artificial intelligence (AI) functions within Translation and Machine Translation (TMT) pedagogy.

By merging the underlying mathematical factor structure with students' explicit behavioral responses, this discussion bridges the gap between statistical groupings and real-world pedagogical dynamics.

5.1 The Latent Dimensions of AI-Facilitated TMT Learning

The PCA optimized the original 10-item instrument into a refined, highly stable 3-factor framework that explains 58.85% of the total variance. Rather than analyzing survey questions in isolation, this factor structure exposes the underlying cognitive, affective, and behavioral frameworks that govern how university students interact with educational AI.

TOTAL VARIANCE EXPLAINED: 58.85%		
Component 1	Component 2	Component 3
Active Growth & Positive Adoption	Dynamic Motivation & Critical Inquiry	Individualization vs. Collaborative Peer Teamwork
(21.83%)	(18.95%)	(18.07%)

The data distinctly groups into three core dimensions:

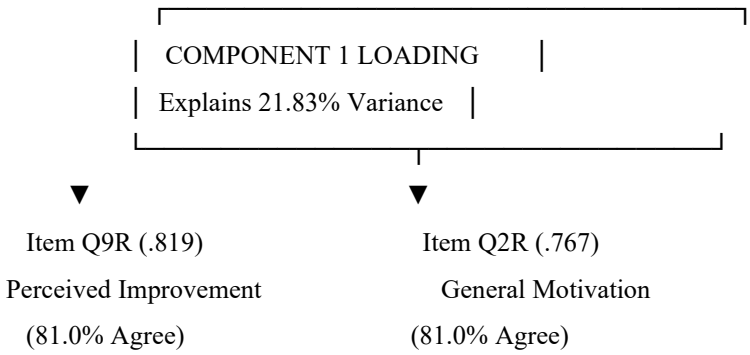
Component 1 (Active Growth and Positive Adoption): Captures the direct link between user satisfaction, engagement, and perceived competence gains.

Component 2 (Dynamic Motivation and Critical Inquiry): Highlights a sophisticated cognitive landscape where intense personal motivation coexists with systemic skepticism.

Component 3 (Individualization vs. Collaborative Teamwork): Reveals the primary pedagogical tension of the AI era—the structural friction between solo human-AI interactions and traditional peer collaboration.

5.2 Component 1: Active Growth and Positive Adoption

Component 1 explains the largest share of post-rotation variance (21.83%) and is strongly defined by Q9R (.819) and Q2R (.767). This statistical alignment reveals that students do not view skill improvement and general motivation as separate outcomes; rather, they form a singular, interconnected dimension of the learning experience.



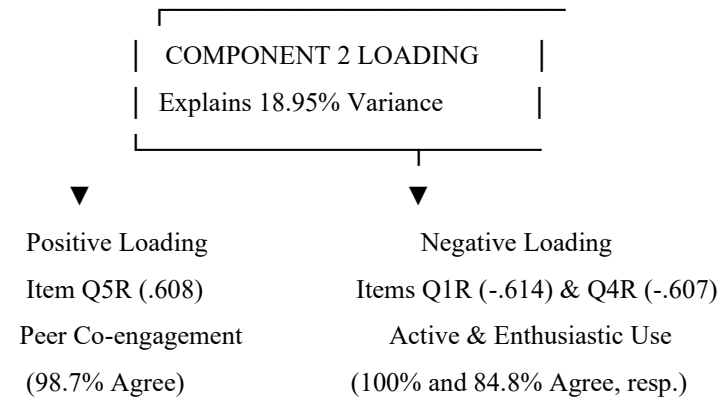
The descriptive statistics strongly support this. A substantial 81.0% of respondents reported clear perceived improvements in executing TMT exercises through AI engagement (Q9R), a figure that mirrors the 81.0% agreement rate for general interest and motivation (Q2R).

This uniform distribution indicates that AI tools provide an immediate linguistic scaffold—offering rapid vocabulary expansions, alternative syntactic structures, and real-time stylistic variations. This immediate validation loop accelerates the transition from abstract translation theory to practical application, directly fueling student confidence and intrinsic morale.

Furthermore, during the item-reduction process, Q10R (active application of AI feedback to improve performance, 98.7% agreement) exhibited a low extraction communality (.355) in the initial 10-item pool, meaning its overwhelming consensus was so absolute that it acted as an ambient, foundational baseline across the entire cohort rather than a variable factor.

5.3 Component 2: Dynamic Motivation and Critical Inquiry

Component 2 (18.95% of variance) presents a highly sophisticated, bipolar cognitive profile. It groups Q5R (.608) positively alongside the negative loadings of Q1R (-.614) and Q4R (-.607).



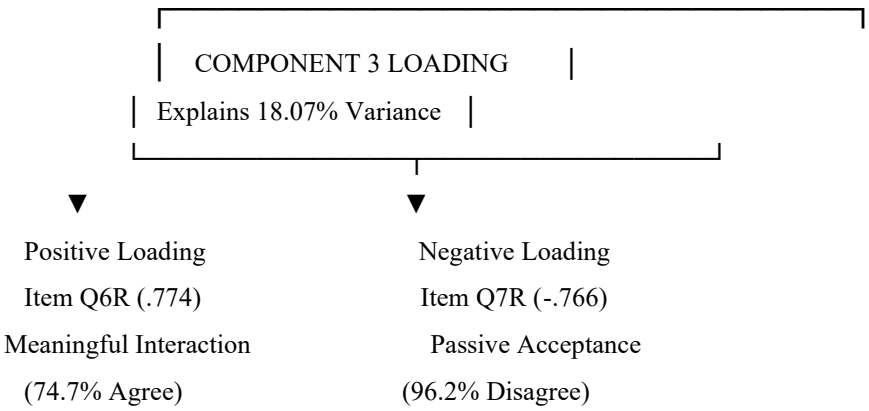
To interpret this matrix accurately alongside the descriptive data—where active participation (Q1R) reached a flawless 100% positive consensus and personal enthusiasm (Q4R) stood at 84.8%—this structural inversion indicates that intense personal engagement operates in a distinct, converse orientation to basic, task-compelled group presence (Q5R, 98.7%).

More importantly, this dimension reflects a student body that is highly motivated precisely because they are cognitively challenged. The descriptive data in Section 5.4 underscores this sophisticated critical posture: 78.4% of students actively verify AI outputs for accuracy (Q8R), while an overwhelming 96.2% reject passive, unthinking text adoption (Q7R).

The negative/positive polarizations within this component prove that students do not treat AI as an effortless shortcut to bypass thinking. Instead, they interact with the technology as an adversarial cognitive partner, using its machine-generated text as a tentative hypothesis that provokes rigorous error analysis, comparative linguistic judgment, and post-editing scrutiny.

5.4 Component 3: Individualization vs. Collaborative Teamwork

Component 3 (18.07% of variance) isolates the most critical pedagogical challenge revealed by the PCA, characterized by a sharp polarization between Q6R (.774) and Q7R (-.766).



While students strongly agree that their coursework structures compel them to operate in groups alongside AI (Q5R at 98.7%), the actual depth of meaningful interaction and teamwork (Q6R) registered the lowest direct agreement rate in the dataset at 74.7%, alongside the highest level of uncertainty (22.8% neutral responses).

The inverse statistical relationship between meaningful peer teamwork (Q6R) and passive text acceptance (Q7R) exposes a major risk of AI integration: the formation of isolated student-AI dyads.

Because large language models provide highly personalized, instantaneous, and responsive feedback loops, they can easily fulfill a student's immediate cognitive needs. This efficiency inadvertently reduces the perceived necessity of consulting classmates.

Consequently, while students are physically sitting in groups or working on the same assignment, their actual cognitive processing becomes highly individualized. Unless collaborative workflows are intentionally designed to require human-to-human debate, peer interaction risks degrading into superficial task-sharing rather than deep, social-constructivist learning.

5.5 Pedagogical Implications and Summary

The empirical evidence soundly refutes old-school assumptions that student responses to GenAI are defined by hesitation, confusion, or passive reliance. Instead, the data reveals a high-consensus, technologically literate cohort capable of balancing enthusiastic adoption with rigorous quality control.

To maximize the value of these findings, translation instructors must shift away from defensive, restrictive policies and embrace AI-assisted workflows.

6. CONCLUSION

To provide the responses to the research questions, we humbly interpret as follows: This study demonstrates that the pedagogical integration of artificial intelligence into Transinterpretation (TMT) instruction can generate meaningful improvements in student motivation, engagement, and perceived competence when implemented within a structured instructional framework. Survey findings reveal that AI does not automatically produce learning gains; rather, its effectiveness depends on deliberate task design, instructor mediation, and editorial refinement. The predominance of “partial” responses across engagement and improvement indicators suggests that AI functions most productively as a cognitive scaffold rather than a substitute for human expertise.

By aligning AI-assisted structural translation exercises with complex grammatical structures, domain-specific vocabulary, and process-oriented evaluation, the study proposes a balanced model that preserves academic rigor while embracing technological innovation. Although limited in scale, the findings provide preliminary empirical support for a human-centered, AI-assisted transinterpretation pedagogy that responds to institutional innovation demands and evolving professional expectations in Vietnamese EFL contexts. The high rates of critical inquiry justify moving away from restrictive, anti-AI policies. Instead, universities should adopt a human-in-the-loop approach where instructors curate complex lexical challenges (such as phrasal verbs and criminal law terminology) that force students to critique and improve AI outputs.

6.1 Limitations and Future Research

This study is limited to a single institutional context and relies partly on self-reported data. Future research should employ longitudinal designs, controlled experimental comparisons, and cross-institutional samples to validate and extend the proposed framework.

6.2 Pedagogical Implications

-For instructors

Instructors should move beyond merely allowing AI use and instead design structured, cognitively demanding tasks that require analysis, justification, and revision of AI outputs. Editorial intervention, structure selection, and domain-specific vocabulary integration are essential to ensure AI supports—rather than replaces—deep transinterpretative competence.

-For Researchers

Researchers should further investigate AI-assisted translation pedagogy through longitudinal, experimental, and cross-institutional studies. Future research should measure actual performance outcomes, not only self-reported perceptions, and examine how instructional mediation influences the relationship between AI use, motivation, and competence development.

-For institutions

Institutions should develop clear policies that balance AI integration with academic integrity, aligning classroom innovation with assessment standards. Investment in professional development, AI literacy training, and curriculum redesign is essential to ensure technology adoption enhances learning outcomes while preserving unaided competence and examination reliability.

-For Students

Students should treat AI as a cognitive scaffold rather than a shortcut. By critically evaluating, questioning, and revising AI-generated translations, learners can strengthen structural awareness, lexical precision, and interpretive reasoning, thereby developing independent competence necessary for examinations and professional translation practice.

References

Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). Routledge.

Beatty, K. (2013). Teaching and researching computer-assisted language learning. Routledge.

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., Saddiqui, S., & Rozenberg, P. (2018). Contract cheating: A survey of Australian university students. *Studies in Higher Education*, *44*(11), 1837–1856.

Brown, et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: ensuring academic integrity in the era of chatgpt. Innovations in Education and Teaching International. https://doi.org/10.1080/14703297.2023.2190148

Dua, B., & Gupta, A. (2026). Developing an integrated AI pedagogical framework for pre-service teachers. *Review of Artificial Intelligence in Education*, *7*(1), e091. https://doi.org/10.37497/rev.artif.intell.educ.v7ii.91

Ellis, R. (2018). Reflections on task-based language teaching. Multilingual Matters; Channel View Publications. https://doi.org/10.2307/jj.22730570

Elshall, A. S., & Badir, A. (2025). Balancing AI-assisted learning and traditional assessment: The FACT assessment in environmental data science education. *Frontiers in Education*, *10*, Article 1596462. https://doi.org/10.3389/feduc.2025.1596462

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Stadler, M. (2023). Chatgpt for good? on opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kiraly, D. (2000). *A social constructivist approach to translator education: Empowerment from theory to practice*. St. Jerome Publishing.

Mimoudi, A., & Mokhtari, K. (2025). 'aia-pcek': a new framework for teaching with AI. Cogent Education, 12(1). https://doi.org/10.1080/2331186x.2025.2563171

Newmark, P. (1988). *A textbook of translation*. Prentice Hall.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Ursachi, T.-M., & Dascalu, M.-I. (2025). Curriculum to immersion: a conceptual framework of artificial intelligence-assisted scenario generation in extended reality for primary and

secondary education. Electronics, 14(24), 4955. https://doi.org/10.3390/electronics14244955

Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.

Wan, Y., Li, R., Li, W., & Du, H. (2025). Impact pathways of ai-supported instruction on learning behaviors, competence development, and academic achievement in engineering education. Sustainability, 17(17), 8059. https://doi.org/10.3390/su17178059

Wei, et al. (2026). *AI-Driven Process-Oriented Translation Instruction to Cultivate Critical Thinking Among Chinese EFL Learners.* International Journal of Web-Based Learning and Teaching Technologies. DOI: 10.4018/IJWLTT.409962.

Zhai, X. (2022). Chatgpt user experience: implications for education. arXiv preprint. https://doi.org/10.2139/ssrn.4312418

#7 Translation-majored Students' ChatGPT Usage Pattern and Translation Self- Efficacy

Quan Ngo[1], Chau Nguyen[2]

Abstract

AI tools such as ChatGPT have radically transformed the teaching and learning processes. While ChatGPT is widely used, its impacts on learners' confidence and capabilities are concerning. Therefore, this research aims to examine the relationship between translation-majored students' ChatGPT usage patterns and their reported translation self-efficacy. Fifty-five students majoring in Translation were involved in a three-part survey which includes questions on demographic information, common ChatGPT usage patterns in doing translation tasks, and the self-efficacy levels adapted from the Translating Self-Efficacy Scale (TSE-C). Differences in translating self-efficacy among groups of users with different ChatGPT usage patterns were analyzed using one-way ANOVA, followed by Games-Howell post-hoc tests for pairwise comparisons. The findings reveal higher confidence among students using ChatGPT to analyze and plan their translation tasks compared to those who used ChatGPT for drafting translations. The study results inform lecturers' instructional design, thus guiding students to effectively use AI tools and develop self-efficacy in translating tasks.

Keywords**:** *ChatGPT, ChatGPT usage patterns, translation self-efficacy*

1. Introduction

AI has become increasingly significant in education with its relevance in language learning and translation. There are numerous AI tools available today; however, ChatGPT has gained much popularity because of its ability to provide responses like human beings and provide quick support for educational activities. In translation learning, students use ChatGPT not only to search for information but also to understand vocabulary, explore cultural meanings, revise their translations, and create translation drafts. According to Ngo (2023), ChatGPT is an advanced natural language processing system that supports users through interactive communication. Therefore, the application of AI-based translation has become more widespread today.

The growing use of ChatGPT entails both benefits and concerns. Firstly, AI software can help learners acquire language skills and improve their efficiency in handling translation tasks (Nguyen et al., 2025). Students may feel less anxious while dealing with complex texts since AI will give them immediate explanations. However, some specialists point out that relying too much on ChatGPT will make students less independent and unable to build their translation skills (Ma, 2024; Nguyen et al., 2025). Instead of analyzing texts and solving the problem themselves, students will become too dependent on computer-generated results, which negatively impacts students' self-efficacy.

In translation pedagogy, students' translating self-efficacy revolves around their confidence in grasping the source text, diagnosing the challenges of the translation, coming up with the solutions to overcome the challenges, and carrying out effective translation. Many

[1] *School of Languages, International University, VNU HCM, Vietnam; Email: ngophamtuongquan@gmail.com*
[2] *Corresponding author, School of Languages, International University, VNU HCM, VietNam; Email: ntnchau@hcmiu.edu.vn*

previous studies concentrated on the significant effects of self-efficacy on the students' motivation, persistence, and performance (Yang et al., 2021; Yuan et al., 2024). Stronger self-efficacy is associated with the students' willingness to face challenges and the ability to work alone. Although numerous studies examined AI-assisted learning and translation competence, few addressed the connecting point between students' utilization of ChatGPT and their translating self-efficacy. As ChatGPT is seemingly morphing to become more of a common part of translation, looking into this relationship is critical for educators and researchers.

Thus, the study explores how students are using ChatGPT for translation and how translating self-efficacy manifests in students with diverse usage patterns. The study employs Borg's (2018) translation process as well as the Human-AI Writing Continuum by Rowland (2023) to analyze students' AI-assisted translation activities.

This study seeks to answer the following research questions:

1. How do students use ChatGPT in translation practices?
2. To what extent do the levels of translating self-efficacy differ across students with different ChatGPT usage?

2. Literature Review

2.1. Translation Self-Efficacy

Bandura (1997) described self-efficacy as self-judgment of the capacity to perform a task that demands one's skill set. Also, these self-beliefs generally play a part in shaping the approach, persistence as well as cognitive and emotional responses to a challenge (Bandura, 1995). Therefore, efficacy beliefs, as well as skills, are required for appropriately performing functions (Bhati & Sethy, 2022). In translation, the term "translation self-efficacy" refers to the confidence of a student in their ability to complete a translation project (Yang et al., 2021). This concept includes the ability to understand, solve problems, make decisions, and produce written output during translation, in addition to the ability to use language effectively.

Research in translation education shows that students with strong self-efficacy are more willing to tackle difficult texts, try out different strategies, and learn from feedback. Self-efficacious people address challenges without much dependence on outside help (Bolaños-Medina & Núñez, 2018). Students who trust their abilities are also more likely to develop self-regulated learning and professional confidence (Liu & Yu, 2019).

Yang et al. (2021) developed the translating self-efficacy scale (TSE-C) - a tool for non-native English speakers, including language learners, translation students, and non-language learners who wish to gain translation skills in preparation for a career as translators. The TSE-C scale is based on Bandura's (1997) self-efficacy framework but is specifically adapted to the actual translation competence required. The TSE-C scale consisted of 15 items grouped into three connected domains: Internal Competence, Psycho-physiological Competence, and External Competence. While Internal Competence (IC) depicts students' confidence in their language proficiency and abilities to solve translation-related problems, the Psycho-physiological Competence (PPC) reflects emotional and physiological factors that impact translation performance like managing stress and self-regulating. The last domain refers to students' confidence in handling external aspects of translation like using resources, managing time and communicating with teachers or clients (Yang et al, 2021).

Translation performance is considered non-linear because it involves many different layers, including contextual factors and individual learner characteristics. Instead of focusing

only on objective test results, the TSE-C scale pays attention to learners' self-concepts and provides valuable insights into the relationship between confidence, motivation, and self-regulation. In this way, it helps explain how learners develop translational competence. Therefore, TSE-C is an appropriately developed scale designed to systematically measure these individual differences.

Self-efficacy also affects how student translators analyze texts, make decisions, revise their translations, and use technology tools, including AI systems such as ChatGPT, during different stages of the translation process. In this context, ChatGPT can support students throughout the translation process and help strengthen their confidence.

2.2. Borg's (2018) Translation Process

Translation can be viewed as a complex task involving multiple stages of translation. In the first stage of the translation process, the translator spends a significant amount of time becoming acquainted with the source text (ST). Acquaintance is achieved through a close reading of the ST, background research, and translation planning. Borg (2018) demonstrated that before the translator even starts on a draft, the translator reads the ST through several times and immerses themselves in the context, using dictionaries and other sources. This is important because studies have shown that some translators do not read the full text in advance (Borg 2018). The goal during this first phase is to make sure the translator fully comprehends the ST.

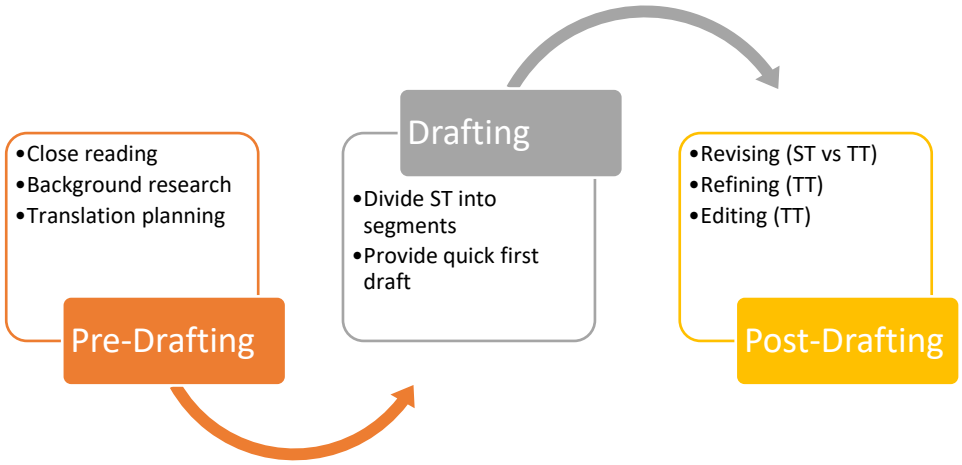
The next stage was the drafting stage. Borg (2018) stated that the translator divides the ST into smaller work units, used only a few external resources, and produces a quick first draft of the TT. Drafting and revising are distinct stages, unlike older approaches to translation, which would have done both simultaneously.

The translator then takes some time away from the draft and begins the last stages which involves: revising, refining, and editing. In the revision one, the translator goes over the text several times, consulting dictionaries and other works and rewriting passages when necessary. This stage involves close comparison of the source and target texts. The next one concentrates on refining the artistic manner of the target text (TT). Within this period, the focus moves toward the target text (TT) instead of the source text (ST). The translation professional improves the artistic manner, selection of words, coherence, spelling, and punctuation. The final one involves careful reading for errors. The translation professional reads the writing with a loud voice so that the rhythm and human flow of the language can be confirmed. Small fixes are performed by the translation professional before the work is sent. Borg (2018) states that translation professionals do not usually return to the source text (ST) during this final stage. A focus on the natural quality and the ease of reading of the target text (TT) is maintained instead.

According to Borg (2018), the number of translating phases is subject to individual translators and may not be limited to a certain number, specifically three phases as commonly confirmed in previous studies (Mossop, 2000). However, in the current study which examines students' ChatGPT use behaviors, the three phases after drafting share the underlying utilization pattern of the AI tool, so they are referred to as Post-Drafting in this study.

Figure 1

Translation Process



2.3. ChatGPT and the Application of ChatGPT in Translation

ChatGPT is based on Generative Pretrained Transformers (GPT; Brown et al., 2020; OpenAI, 2023) and trained using the Reinforcement Learning from Human Feedback (RLHF). As a result, it can generate natural responses, answer questions, and support ongoing interaction. With a large amount of human-written data, such as books, articles, reviews, and online discussions, the model was trained to recognize language patterns and respond to different topics more humanlike, thus being widely used in language-related tasks, including translation (Wang et al., 2023). With ChatGPT's assistance, students can explore difficult source texts before drafting, generate translation suggestions, and identify errors of the translations (Brahmana et al., 2025; Nguyen 2024). ChatGPT was also recognized as a useful tool for students to revise their translated text by prompting the AI tool to provide feedback on their translations (Yang et al., 2025) or to create alternative versions to evaluate different styles of translation (Nguyen, 2024).

The level of AI involvement in translating, however, varies among students. Some mainly use it for background research or improving language quality, while others rely on it more heavily to complete translation tasks. This also leads to the problem of students relying too much on ChatGPT to perform the entire translation process, completely replacing humans, instead of just using it as a support tool. Scholars caution that AI-assisted workflows risk diminishing translators' autonomous decision-making if used passively or uncritically (Pym, 2014). As a result, examining students' AI usage patterns requires a comprehensive model that depicts students' different levels of AI involvement in doing translation tasks.

2.3 The Human-AI Writing Continuum

Rowland (2023) proposed the Human-AI Writing Continuum to model the relationships between human writers and AI tools in writing. This continuum provided a basis to establish the Human-AI Translating Continuum, which is the continuum used in this research. At one end is human-led where no AI is used (level 1), while at the other end is AI-led, where most of the text is generated by AI (level 9).

Between these two endpoints exist several types of human-AI interaction. They include writing guided by AI, wherein the artificial intelligence assists generating ideas; writing assisted by AI, in which AI helps construct sentences and provide linguistic input; and

collaborative writing, where human beings and AI work together during the writing process. This model underscores the nonlinear and highly contextual nature of human-AI interaction.

The continuum model is significant because it illustrates that the association between man and machine is dynamic and highly situational. For this study, the model would be used to determine how students use ChatGPT when translating. For instance, students could use ChatGPT to look up cultural words and concepts but become more independent during drafting and revising. By placing students' behaviors along the continuum, the study can better understand patterns of AI use, levels of dependence, collaboration strategies, and their possible influence on translation self-efficacy.

2.4. Research gaps

Previous works have explored mostly machine translation systems and CAT tools but have rarely looked at conversational AI like ChatGPT and its relationship with translation learning and self-perceptions. Similarly, research on the usage of ChatGPT in education has mostly covered academic writing, rather than translation learning. Accordingly, further research is still necessary to uncover how student translators are using ChatGPT and how this usage relates to their translation self-efficacy. The present study thus intends to examine the relationship between ChatGPT use and translation self-efficacy in student translators.

2.5. Conceptual Framework

The Rowland's (2023) Human-AI continuum and Borg's (2018) translating process are adopted to depict the dynamic roles of ChatGPT and human in translation tasks. Instead of treating AI use as a simple "use or not use" issue, the continuum presents a range of interactions in which human contribution and AI assistance vary throughout the translation process. Based on nine levels of Rowland's continuum, the current study's seven levels were adopted to represent different patterns of students' ChatGPT usage during translation tasks. Because translating tasks do not require students to generate new ideas, level 4 of the original continuum was adapted into "Students ask ChatGPT to analyze the assignment topic, explore background information or cultural context, clarify difficult terms, idioms, or technical vocabulary before translating. However, the student translates and reviews the final version." The nature of the level was maintained with AI assistance in content analysis and human's control of the translating/writing task. The original levels 6 and 8 were omitted as they are closely related to writing tasks and may not reflect the nature of translation. Specifically, paraphrasing ideas from other sources and generating external texts, e.g. an abstract for student-written paper, are not relevant to translating tasks, thus being removed. The seven remaining levels are adapted and mapped with the proposed translating process as demonstrated in table 1.

Table 1

Levels of Interaction with AI in Translation Phases

Level	Description	Translating Phase wherein AI is used
1	Students perform all tasks independently, including reading and analyzing the source text, making lexical and stylistic decisions, drafting and revising the target text, and completing quality checks without AI input.	None
2	As per level 1, but students use ChatGPT to correct grammar, vocabulary, or spelling.	Post-Drafting

3	As per level 1, but ChatGPT is asked for feedback on how the student translation can be improved. Students adopt or adapt recommended changes.	Post-Drafting
4	Students ask ChatGPT to analyze the assignment topic, explore background information or cultural context, clarify difficult terms, idioms, or technical vocabulary before translating. However, the student translates and reviews the final version.	Pre-drafting
5	Students use ChatGPT to translate the whole text, then they review the output and edit it by themselves.	Drafting
6	ChatGPT generates the first draft translation. After that, the student has a "conversation" with ChatGPT until they can achieve the best version.	Drafting
7	Students adopt without change, and ChatGPT is in the role of translating the whole translation task.	All

Combining these two frameworks allows researchers to examine both the stages in which students use ChatGPT and the extent of AI involvement during translation tasks. The Pre-Drafting stage aligns with Level 4 of the Human-AI Continuum, in which students apply ChatGPT to get a summary, search for backgrounds, and promote their comprehension, and they translate texts on their own. In this stage, AI works as a tool of support instead of substitution. The Drafting stage includes Levels 5 and 6, where ChatGPT plays a larger role in generating translations. Students revise and edit the AI-generated output based on their own linguistic competence. At Level 6, students repeatedly prompt ChatGPT to revise translations. The last translation phase corresponds to Level 2 and 3, in which students mainly adopt ChatGPT to provide feedback and suggestions. These typologies help explain students' behaviors of using ChatGPT and facilitate in comparing translation self-efficacy differences across different AI user groups.

3. Methodology

3.1. Study design and Participants

This study employs a quantitative design to explore the relationship between translation-majored students' ChatGPT usage in translating texts and their translation self-efficacy. Data analysis was conducted using descriptive statistics and one-way ANOVA analysis in Jamovi – an open statistical software. The research is divided into two parts: the first part deals with how students used ChatGPT when performing translation tasks, and the second part explores how students rated their confidence when performing the translation task without ChatGPT assistance.

The sample of this study includes 55 students selected from the population of English-majored students from a university in Vietnam. However, two responses were removed for two reasons: (1) one student reported no AI use in their translating practice, so this response is not related to the research questions; and (2) according to standard course sequencing and prerequisite requirement, one student response shows a mis-match between the year of study and the taken translation courses, so their answer is considered not reliable. As a result, there are a total of 53 participants. All of them have or have been taking translation courses as part of their academic program, ensuring a shared foundation in translation theory and practice.

3.2. Instrument:

A Google Forms survey was used to gather data on students' ChatGPT usage and translation self-efficacy. The survey with the three distinct parts completely addressed the objectives of the study.

The first section contains questions about participants' background and demographics, including students' year of study, the number of translation courses they had taken, and their self-assessed English proficiency based on the CEFR framework.

Then, students selected their usage pattern, or the most suitable way of using ChatGPT to translate texts, among seven predetermined options following the Human-AI Translating Continuum, ranging from not using ChatGTP to completely relying on this AI tool.

The third section adapted Yang et al.'s (2021) Translating Self-Efficacy Scale with Likert-scale items that evaluated Internal, Psycho-physiological, and External Competence.

The validity and reliability of the instrument was strengthened by implementing a pilot study and adapting Yang et al. (2021)'s validated scale.

Section 2 of the questionnaire were adapted based on the emergent frameworks of Borg's Translation Process and Rowland's (2023) Human-AI writing continuum. A pilot study with 12 senior students who have participated in compulsory and elective translation courses was conducted. Their examples of ChatGPT prompts were used to identify any ChatGPT usage pattern which was not included in the section 2 of the questionnaire. For example, participant 9 of the pilot study explained their ChatGPT usage as follows.

I paste the whole text I already translated as well as the source textand ask Chat GPT "Please evaluate my translation compared to this source text following the rubrics: Grammar Accuracy, Vocabulary Use, and How compatible of the two texts. Also, please provide some improvements related to cultural differences for my translation, clafifying that whether the target language can fully interpret the source language but the target readers can still understand what the translation is about."

The explanation matches with option 3 in the list of 7 ChatGPT usage pattern which states "I do the task by myself, then use ChatGPT to give feedback on the quality of translation, and I adapt the recommended changes". The pilot study did not reveal any using behavior that cannot be matched with the 7 predetermined usage patterns of the questionnaire, which confirms the validity of section 2.

Regarding section 3, the scale was mainly adapted from Yang et al.'s (2021) Translating Self-Efficacy Scale. Before Yang et al. (2021) developed the TSE-C scale, there were a lot of preliminary scales measuring students' translating confidence in translation, such as the scales developed by Stansfield et al. (1992), Bolaños-Medina and Núñez (2018), and Konttinen (2021). However, each of these scales had certain limitations. Stansfield et al. (1992) introduced what was considered the first "translation ability self-assessment questionnaire" (as cited in Abdel Latif & Alhamad, 2025). The scale mainly measured students' perceived ability to complete translation tasks, but it did not include important translation competencies. Another limitation was the lack of validation data in the published study. Later, Bolaños-Medina and Núñez (2018) developed the first standardized translation self-efficacy scale. However, their research focused mainly on English–Chinese translation and used a relatively small sample size. Konttinen's (2021) scale was developed more recently, but it focused on students' self-efficacy in translation company simulations rather than translation practice in educational settings. Taking all the above considerations and gaps of previous translating self-efficacy, the scale by Yang et al. developed in 2021 met all the requirements. The translating self-efficacy

scale (TSE-C) is a valid and dependable tool for non-native English speakers, including language learners, translation students, and non-language learners who wish to gain translation skills in preparation for a career as translators. The TSE-C scale is based on Bandura's (1997) self-efficacy framework but is specifically adapted to the actual translation competence required. It views translation self-efficacy as something complex and multilayered that affects students' level of confidence in their ability to translate well.

To ensure the reliability of the Translating self-efficacy scale, reversed items were added into the last construct, External Competence, in order to test if participants pay sufficient attention to reading items and choosing responses. Cronbach's alpha was measured, and items were removed to ensure the internal consistency.

Efficacy in…	Cronbach's Alpha	Removed Items	Updated Cronbach's Alpha
Internal Competence (IC)	.773	IC2	.805
Psycho-physiological Competence (PC)	.562	PC5	.810
External Competence (EC)	.610	EC4	.667

The final version of Translation self-efficacy scale, therefore, includes 12 items with 6 items measuring efficacy in Internal Competence, 2 items for Psycho-physiological Competence, and 4 items for External Competence.

3.3. Data analysis

The data were analyzed using Jamovi to examine differences in translation self-efficacy among students who used ChatGPT for translation in three different stages: Pre-Drafting, Drafting, and Post-Drafting. Descriptive statistics including the frequency and percentage of Agree and Strongly Agree options were reported to provide an overview of students' confidence when translating without using ChatGPT.

A one-way ANOVA was then conducted to compare translation self-efficacy across three groups of ChatGPT usage. In Jamovi, the dependent variable was labelled as TSE (Translating Self-Efficacy score), which is the sum of participants' answers for 12 items adapted from Yang et al. (2021). The independent variable was labeled as ChatGPT_Stage and consisted of three groups formed from seven predefined usage patterns based on the Human–AI Translating Continuum and Borg's (2018) Translation Process model.

4. Results

4.1. Descriptive Analysis of Variables

Most participants (53 out of 54) employed ChatGPT to assist their translation process; however, the data indicates that students employed ChatGPT in vastly different ways when working on their translation tasks. No students used ChatGPT to do the whole translation without any human input, while others utilized the AI tool in different stages.

Table 2

ChatGPT Usage Pattern

ChatGPT is used in	Frequency	Percentage
Pre-Drafting	22	41.51%
Drafting	20	37.74%
Post_Drafting	11	20.75%

The two largest groups of students used ChatGPT for Pre-Drafting and Drafting phases, with 41.51% and 37.74% of participants, respectively. Most students used ChatGPT in the Pre-Drafting phase to explore background information, cultural context, or difficult vocabulary before translating the text independently. A significant part of students (20.75%) employed ChatGPT for revising and editing while producing the draft on their own. However, a large proportion of participants used ChatGPT to generate draft translations and then reviewed as well as edited the output themselves or with ChatGPT assistance. Overall, the finding suggests that most students viewed ChatGPT as a support tool rather than a replacement for translation. In other words, a more balanced and supportive use of ChatGPT was frequently employed rather than using generative AI tools exclusively or minimally.

Table 3

ChatGPT Usage Pattern by English Proficiency

Proficiency	Total Number	Pre-Drafting	Drafting	Post_Drafting
B1	5	40.00%	40.00%	20.00%
B2	17	29.41%	41.18%	29.41%
C1	29	51.72%	34.48%	13.79%
C2	2	0.00%	50.00%	50.00%

As the number of participants with B1 annd C2 levels are relatively limited compared to the other two, their responses should be considered with caution. Regarding the two largest groups of participants, the ones with higher proficiency (C1) has a larger proportion of students using ChatGPT for the Pre-Drafting stage while the B2 group tend to rely on ChatGPT for Drafting.

Table 4

ChatGPT Usage Pattern by Translation Experience (via taken translation courses)

Translation Experience	Total Number	Pre-Drafting	Drafting	Post-Drafting
COMPULSORY	38	50.00%	36.84%	13.16%
COMPULSORY_ELECTIVE	15	20.00%	40.00%	40.00%

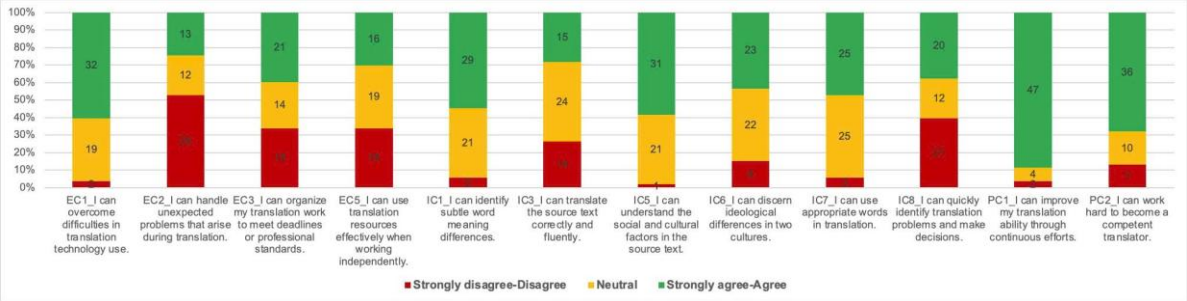
There are significant differences between the utilization of ChatGPT in the Pre-Drafting and Post-Drafting of the two groups. While those with lower exposure to translation courses

(only taking compulsory translation courses) mainly used ChatGPT for Pre-Drafting, the more experienced group tend to depend on ChatGPT for Drafting and Post-Drafting. As Post-Drafting is the stage with limited ChatGPT use compared to the other two, this remarkable tendency among the experienced group gives an insightful perspective on students' ChatGPT using behavior.

Participants' self-reported Translating Self-Efficacy (TSE) varied across dimensions, including Internal Competence (IC), Psycho-physiological Competence (PC), and External Competence (EC). Reversed items have been reworded for consistent reporting.

Figure 2

Frequency of TSE Responses

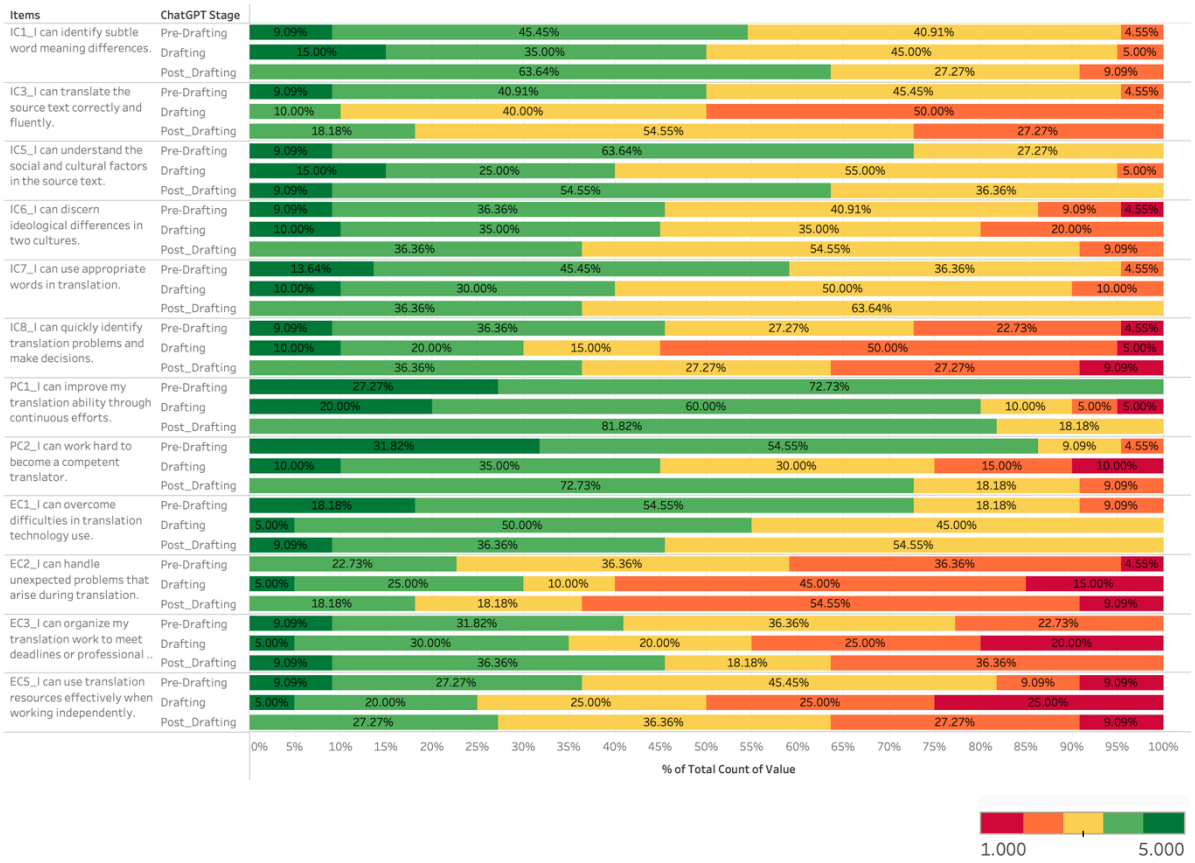


Overall, the results indicate that students reported relatively high levels of translation self-efficacy (TSE), as most responses were concentrated in the "agree" and "strongly agree" categories across items. Higher levels of confidence were particularly evident in items related to translation improvement and self-assessment (PC1, PC2, EC1), where a large majority of respondents selected positive response options. However, confidence was comparatively lower when encountering external challenges (EC2, EC3, EC5) or tasks requiring cognitively demanding skills such as translating source texts correctly and fluently (IC3), as well as identifying and solving translation problems (IC8).

The TSE responses also followed divergent patterns across three ChatGPT usage groups, namely those use ChatGPT in the Pre-Drafting, Drafting and Post-Drafting stages.

Figure 3

Frequency of TSE Responses by ChatGPT Usage Groups



Users of ChatGPT for Pre-Drafting demonstrated the strongest confidence levels. For most items, they consistently exhibit the largest percentages of positive responses and lowest percentages of negative evaluations, whereas Drafting and Post-Drafting users displayed more varied response patterns. Negative responses are limited across ChatGPT usage groups except for some special cases. Specifically, Drafting and Post-Drafting groups' answers to EC2 (I can handle unexpected problems that arise during translation), Drafting groups' answers to EC5 (I can use translation resources effectively when working independently), IC3 (I can translate the source text correctly and fluently) and IC8 (I can quickly identify translation problems and make decisions) include a substantial amount of negative choices (more than 50%). More than a third of Post-Drafting participants also reported low confidence in EC3, EC5, and IC8. Nevertheless, in general, positive perceptions dominated across all groups, suggesting that students generally felt capable of performing translation tasks regardless of how they incorporated ChatGPT into their translation process. Their total TSE scores – calculated by adding up scores for 12 items – can be summarized as follows.

Table 5

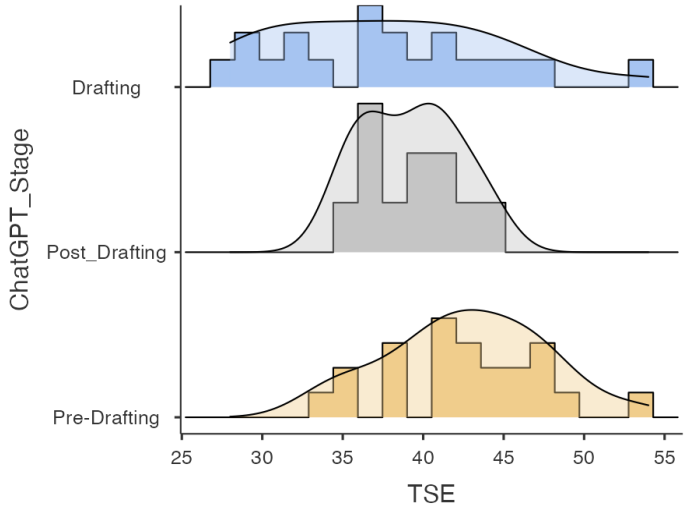
TSE Scores across ChatGPT Usage Groups

	ChatGPT_Stage	N	Missing	Mean	Median	SD	Minimum	Maximum
TSE	Pre-Drafting	22	0	42.7	43.0	4.84	34	53
	Drafting	20	0	37.8	37.0	7.01	28	54

ChatGPT_Stage	N	Missing	Mean	Median	SD	Minimum	Maximum
Post_Drafting	11	0	39.2	40	2.99	35	44

Figure 4

TSE Scores Distribution across ChatGPT Usage Groups



4.2 One-way ANOVA Results

Based on the ChatGPT usage patterns, a one-way analysis of variance (ANOVA) was conducted with the dependent variable being the translating self-efficacy score (TSE) and the independent variable being the 3 groups of ChatGPT usage patterns (Pre-Drafting, Drafting, Post-Drafting). First, Levene's and Shapiro-Wilk tests were conducted to test the assumption of homogeneity of variance and normality across groups.

Table 6

Assumption Checks

Normality Test (Shapiro-Wilk)

	W	*p*
TSE	0.979	.462

Homogeneity of Variances Test (Levene's)

	F	df1	df2	p
TSE	4.14	2	50	.022

Shapiro-Wilk were not statistically significant (p = .462 >.05), indicating normally distrubted data. However, Levene's test is below the .05 threshold, revealing violated homogeneity of variance (p = .022). As a result, Welch's ANOVA and Games-Howell were used for the data analysis instead of Fisher's and Tukey's HSD. The results of the ANOVA analysis are presented in Table 5 below:

Table 7

ANOVA Results for Differences in Translating Self-Efficacy Across ChatGPT Usage Levels

One-Way ANOVA (Welch's)

	F	df1	df2	p
TSE	4.60	2	31.8	.018

The results showed the statistically significant difference in translating self-efficacy levels between the groups using ChatGPT (p = .018). This suggests that differences in translating self-efficacy levels are associated with different levels of students' use of ChatGPT.

Table 8

ANOVA Post-hoc Comparison of Translating Self-Efficacy Across ChatGPT Usage Levels

Games-Howell Post-Hoc Test – TSE

		Drafting	Post_Drafting	Pre-Drafting
Drafting	Mean difference	—	-1.38	-4.88
	p-value	—	.728	.036
Post_Drafting	Mean difference		—	-3.50
	p-value		—	.042
Pre-Drafting	Mean difference			—
	p-value			—

Games-Howell post-hoc analysis showed a statistically significant difference in translating self-efficacy between Pre-Drafting and Drafting (p = .036), with students using ChatGPT in the Pre-Drafting phase reporting higher translation self-efficacy. The choices of using of ChatGPT in the Pre-Drafting or the Post-Drafting phases also correlate with the different levels of translating self-efficacy (p = .042). In both cases, students using ChatGPT in Pre-Drafting reported higher level of confidence compared to those employing ChatGPT for the Drafting and Post-Drafting phases. The TSE scores of Drafting and Post-Drafting showed no significant differences (p = .728 > .05).

Further ANOVA tests were run with independent variable being the proficiency and the taken translation courses respectively. Both tests reveal insignificant statistical differences in TSE among these groups. The results are presented as follows.

Table 9

ANOVA Post-hoc Comparison of Translating Self-Efficacy Across Proficiency Levels

Tukey Post-Hoc Test – TSE

		B1	B2	C1	C2
B1	Mean difference	—	-0.741	-2.59	-6.80
	p-value	—	.994	.796	.511
B2	Mean difference		—	-1.85	-6.06
	p-value		—	.729	.514
C1	Mean difference			—	-4.21
	p-value			—	.759
C2	Mean difference				—
	p-value				—

Table 10

ANOVA Post-hoc Comparison of Translating Self-Efficacy Across Translation Experience (via Translation courses taken)

Tukey Post-Hoc Test – TSE

Taken translation courses that are		COMPULSORY	COMPULSORY & ELECTIVE
COMPULSORY	Mean difference	—	-0.121
	p-value	—	.947
COMPULSORY_ELECTIVE	Mean difference		—
	p-value		—

5. Discussion

The findings provide insights into how students use ChatGPT during translation tasks and how different usage patterns relate to translation self-efficacy.

5.1. Students Use ChatGPT to Support Their Translation Skills

From the results presented, the students used ChatGPT at different stages of the translation process. The most common stage that was used is the Pre-Drafting stage, which accounts for 41.54% of the total number of students. At this stage, it is mainly used by the students to gather background knowledge, cultural knowledge, or even unknown words before creating their translations. The use of ChatGPT for Drafting stage was examined in 37.74% of the participants, who employed ChatGPT to produce draft translations and then refined the products through either repeated interaction with the AI tools or on their own. Meanwhile, 20.75% of students mainly used ChatGPT during the polishing, fine-tuning, and proofreading stages, utilizing the tool to check errors and provide feedback after students completed their translations independently. Within this general trend, participants with different proficiency and translation experience demonstrate diverse tendencies in ChatGPT using behavior. Specifically, whie the higher proficiency group mostly use ChatGPT for Pre-Drafting, the lower one rely on the AI tool for Drafting. This is understandable as students with higher English proficiency face less challenges in language use and start focusing on ensuring the accuracy, naturalness and appropriateness in their translation. As a result, they prefer more ChatGPT assistance for gathering background and cultural knowledge before translating texts. ChatGPT utilization of participants with different translation experiences also yields valuable insights. The group with lower translation experience mainly use ChatGPT for Pre-Drafting and Drafting. On the other hand, the higher experience group favor the use of AI support in the Drafting and Post-Drafting. This illustrates the students' perception of the translator role. While those with fewer experience consider themselves learners and use AI tools to support the process of enriching knowledge and skills, the one taking more translation courses may have familiarized themselves with the translation process and start assigning tasks to their AI assistance.

Unlike previous studies such as Kaiyi (2025) and Dinh (2025), which identified the revision stage as the most common stage of AI use, this study found that students mainly used ChatGPT during the preparation stages. This suggests that students regarded ChatGPT more as a tool for exploring and organizing knowledge than as a complete translation assistant. These differences may be due to differences in students' backgrounds, language abilities, and learning goals. Overall, the findings reveal ChatGPT's role as a support tool throughout the translation process rather than a replacement for human translators.

5.2. Levels of Translation Self-Efficacy and Differences among Groups of ChatGPT Usage

The descriptive analysis showed that students' translation self-efficacy was overall high in all three dimensions: Internal Competence, Psycho-physiological Competence, and External Competence. Among these three competences, students reported higher confidence in Psycho-physiological Competence and lower in External as well as Internal Competence. Based on this information, students had higher confidence in controlling their efforts, persistence, and emotions rather than in their language, decision-making abilities, and external conditions. The low scores for Internal Competence suggest that students still had difficulty analyzing the language, constructing their own meanings, and solving translation problems. Such a conclusion aligns with the results of Dinh's (2025) study, which found that translators experience difficulties when developing their translation strategies and defending their translation choices.

The one-way ANOVA results also revealed a statistically significant difference in translating self-efficacy among groups with different levels of ChatGPT usage ($p = .018$). Post-hoc analysis showed that this difference was primarily between Pre-Drafting and the other two stages. Students who relied more heavily on ChatGPT for generating translations demonstrated

lower confidence in their own translation abilities than students who used the tool more selectively. Additionally, there are no statistically significant differences between students with different English proficiency and translation experience. The findings suggest that proficiency and experience are not significant factors impacting translation self-efficacy, and excessive reliance on AI may reduce opportunities for independent practice and mastery experiences, which Bandura (1997) identified as an important source of self-efficacy development. The results also echo Zhang and Xu's (2025) paradox of self-efficacy which states leveraging AI can boost students' confidence but at the same time intensifying their AI reliance. As a result, an appropriate level of AI integration in teaching and learning translation is required to ease the negative impact of the paradox. Overall, the findings indicate that careful and critical use of ChatGPT may support translating self-efficacy, while overdependence on AI tools may weaken students' confidence in their own translation competence

6. Conclusion

The present study investigated the relationship between the use of ChatGPT and translation self-efficacy among university student translators. Findings from the TSE-C scale (Yang et al., 2021) and one-way ANOVA analysis showed that different levels of ChatGPT use were linked to differences in students' translating self-efficacy. In general, students who employed ChatGPT strategically and moderately had higher confidence in translation tasks compared to those who heavily relied on AI-generated translations.

The findings also showed that students used ChatGPT in different stages of the translation process. For most participants, they made use of the platform in the Pre-Drafting stage, where they would carry out research about the text, clarify terms, and plan on the best approach to take before embarking on the translation process. The remaining participants resorted to the platform to create translations or for proofreading purposes after creating their own versions. However, different levels of proficiency and translation experience can affect their ChatGPT using behavior.

Regarding translating self-efficacy, students reported moderate levels across the three dimensions of internal competence, psycho-physiological competence, and external competence. Among these, psycho-physiological competence attained the maximum rating because the students felt more competent at putting in effort, persisting, and regulating emotions while undertaking translation. Internal competence, on the other hand, scored the least since the students had problems with linguistic analysis and translation strategy making.

Importantly, the study found significant differences in translating self-efficacy among groups with different ChatGPT usage patterns. The ones who extensively used ChatGPT when translating were less likely to display self-efficacy, while those using the tool moderately, such as for generating ideas and looking up terms, showed more self-confidence. It is possible to infer from the results that moderation in the use of ChatGPT could foster self-efficacy in learners.

The findings also provide several implications for translation education. Teachers should encourage selective and goal-oriented use of ChatGPT rather than allowing students to depend on AI throughout the entire translation process. ChatGPT can be effectively used for brainstorming, terminology clarification, and culture lookup while students remain responsible for the main translation decisions. Universities may also consider providing AI literacy training, including prompt writing, evaluation of AI-generated content, and strategies for revising AI-supported translations. In addition, the study contributes theoretically by combining Borg's (2018) Five-Phase Translation Model with Rowland's (2023) Human–AI

Continuum to better explain the relationship between AI involvement and learner autonomy in translation tasks.

Despite these contributions, the study has several limitations. First, the sample size was relatively small and limited to students from one university, which may reduce the generalizability of the findings. Additionally, because the study relied on self-reported survey data, participants may have underreported their AI use or overestimated their self-efficacy levels.

Overall, the study suggests that ChatGPT can serve as an effective support tool in translation education when used carefully and critically. However, passive or excessive reliance on AI may weaken students' confidence and independence in translation tasks. Therefore, maintaining a balance between AI support and human control is important for developing both translation competence and translating self-efficacy in technology-enhanced learning environments.

References

Abdel Latif, M. M., & Alhamad, M. M. (2025). Translation and interpreting self-efficacy: A review of published measures. *Frontiers in Psychology*, 16. https://doi.org/10.3389/fpsyg.2025.1632366

Bandura, A. (1995). *Self-efficacy in Changing Societies*. University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control* (Vol. 11). Freeman.

Bhati, K., & Sethy, T. (2022). Self-efficacy: Theory to educational practice. The International Journal of Indian Psychology, 10(1), 1123-1128. DOI: 10.25215/1001.112

Bolaños-Medina, A., & Núñez, J. L. (2018). A preliminary scale for assessing translators' self-efficacy. *Across Languages and Cultures, 19*(1), 53–78. DOI:10.1556/084.2018.19.1.3

Borg, C. (2018). *The phases of the translation process: Are they always three?* In *Junior College Multi-Disciplinary Conference: Research, Practice and Collaboration—Breaking Barriers* (pp. 79–92). University of Malta. https://www.um.edu.mt/library/oar/handle/123456789/38130

Brahmana, C. R., Sofyan, R., & Mono, U. (2025). Evaluating the strengths and limitations of CHATGPT-generated translations in academic post-editing workflows. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching, 9*(1), 135–151. https://doi.org/10.30743/ll.v9i1.10551

Brown, Tom & Mann, Benjamin & Ryder, Nick & Subbiah, Melanie & Kaplan, Jared & Dhariwal, Prafulla & Neelakantan, Arvind & Shyam, Pranav & Sastry, Girish & Askell, Amanda & Agarwal, Sandhini & Herbert-Voss, Ariel & Krueger, Gretchen & Henighan, Tom & Child, Rewon & Ramesh, Aditya & Ziegler, Daniel & Wu, Jeffrey & Winter, Clemens & Amodei, Dario. (2020). Language Models are Few-Shot Learners. https://doi.org/10.48550/arXiv.2005.14165

Dinh, C. T. (2025). EFL students' perspectives on chatgpt in translation: Exploring AI assistance, motivation, and learning outcomes. *Electronic Journal of E-Learning, 23*(2), 99–116. https://doi.org/10.34190/ejel.23.2.4006

Kaiyi, D. (2025). Research on post-editing strategies for Chinese-english translation based on Chatgpt. *Lecture Notes on Language and Literature, 8*(4). https://doi.org/10.23977/langl.2025.080419

Konttinen, K. (2021). A self-efficacy scale for measuring student progress in Translation Company simulations. *Across Languages and Cultures, 22*(1), 64–81. https://doi.org/10.1556/084.2021.00004

Liu, C., & Yu, C. (2019). Understanding students' motivation in translation learning: A case study from the self-concept perspective. *Asian-Pacific Journal of Second and Foreign Language Education, 4*(1). https://doi.org/10.1186/s40862-019-0066-6

Ma, Y. (2024). The impacts and challenges of CHATGPT on translation teaching in universities. *Journal of Education, Humanities and Social Sciences, 38*, 155–161. https://doi.org/10.54097/fjp9tb72

Mossop, B. (2000). The workplace procedures of professional translators. In A. Chesterman, N. Gallardo San Salvador, & Y. Gambier (Eds.), *Translation in context: Selected*

papers from the EST Congress, Granada 1998 (pp. 39–48). John Benjamins. https://doi.org/10.1075/btl.39.07mos

Ngo, T.T.A. (2023) The Perception by University Students of the Use of ChatGPT in Education. *International Journal of Emerging Technologies in Learning (iJET), 18*, 4-19. https://doi.org/10.3991/ijet.v18i17.39019

Nguyen, H. (2024). Integrating ChatGPT in Translation Learning.In *Annual SEAAIR Conference Proceedings, 4*, 602-611. https://www.researchgate.net/publication/391425619_Integrating_ChatGPT_in_Translation_Learning

Nguyen, T. N. N., Tran, T. T., Nguyen, N. H. A., Lam, H. P., Nguyen, H. M. S., & Tran, N. A. T. (2025). The benefits and challenges of AI translation tools in translation education at the tertiary level: A systematic review. *International Journal of TESOL & Education*, *5*(2), 132–148. https://doi.org/10.54855/ijte.25527.

OpenAI (2023). ChatGPT (Mar 14 Version) [Large Language Model]. https://chat.openai.com/chat

Pym, A. (2014). Translation skill-sets in a machine-translation age. *Meta, 58*(3), 487–503. https://doi.org/10.7202/1025047ar

Rowland, D. R. (2023). *A conceptual framework for discussing the human-AI writing continuum*. ResearchGate. https://doi.org/10.13140/RG.2.2.36445.59361

Stansfield, C. W., Scott, M. L., & Kenyon, D. M. (1992). The measurement of translation ability. *The Modern Language Journal, 76*(4), 455. https://www.researchgate.net/publication/263588437_The_Measurement_of_Translation_Ability

Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D., Shi, S., & Tu, Z. (2023). Document-level machine translation with large language models. In H. Bouamor, J. Pino, & K. Bali (Eds.), *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 16646–16661). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.1036

Yang, C., Chen, J., Hou, S., Wider, W., & Yi, W. L. (2025). How AI Benefits Student Translators: An Exploratory Study on the Impact of ChatGPT Feedback on Translation Proficiency. *Forum for Linguistic Studies, 7*(10), 526–541. https://doi.org/10.30564/fls.v7i10.11080

Yang, Y., Cao, X., & Huo, X. (2021). The Psychometric Properties of Translating Self-Efficacy Belief: Perspectives from Chinese learners of translation. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.642566

Yuan, Y., Wu, J., & Wu, Y. (2024). The influence of translation majors' self-efficacy on their interpreting ability and coping strategies under the guidance of gile's model. *Advances in Social Science, Education and Humanities Research*, 420–435. https://share.google/vHPJvKMBBZxPtdnyF

Zhang, L., & Xu, J. (2025). The paradox of self-efficacy and technological dependence: Unraveling generative AI's impact on University Students' task completion. *The Internet and Higher Education, 65*, 100978. https://doi.org/10.1016/j.iheduc.2024.100978

Bionote

Quan Ngo is a fresh graduate from the School of Languages, International University, Vietnam National University HCMC. Her research interest is the integration of AI in translation and English teaching.

Email: ngophamtuongquan@gmail.com

Chau Nguyen is a lecturer in the School of Languages, International University, Vietnam National University HCMC. She has been teaching Intensive and Academic English courses for more than 10 years. Her current research interest includes educational technology and TESOL methodology.

Email: ntnchau@hcmiu.edu.vn

#15 From Classroom to Real Life: First-Year Students' English-Speaking Use across Learning Contexts

Nguyen Thi Ngoc Chinh[1]

Abstract

The present study aims to examine the English-speaking practices inside and outside the classroom by first-year Vietnamese university students during their preparatory English program in a non-English-speaking environment. The study sample consisted of 50 lower-intermediate EFL learners at the outset of the program, corresponding to levels A2-B1 on the CEFR scale. In the Vietnamese context, classroom-based English speaking is considered essential for developing oral communication skills, while outside-school real-life use is limited, especially for young university students. This study also sought to examine the relationship between in-class speaking and out-of-school English use. The study employed the explanatory sequential mixed-methods research design. Quantitative data collection methods were employed to address the research questions, including the administration of a structured questionnaire to assess the frequency of English-speaking practices among participants. Additionally, semi-structured interviews were used to gain more insights into the participants' out-of-school use of English. Quantitative data were analyzed through statistical methods, while qualitative data were analyzed through thematic analysis. The findings indicate that the students spoke more English inside the classroom than outside the classroom. Inside the classroom, the students spoke English mainly through structured activities, such as answering teachers' questions, discussing in pairs or groups. On the other hand, the participants rarely used English outside the classroom to communicate in real-life situations. The findings reveal a substantial discrepancy between instructional speaking participation and spontaneous English use in authentic communicative contexts.

***Keywords*:** English speaking, EFL context, first-year university students, classroom and real-life communication, mixed-methods study

1. Introduction

English speaking competence is widely recognized as an essential learning outcome (Richards, 2008; Bygate, 1987) in the context of tertiary English language learning and teaching, especially in the context of English as a Foreign Language (EFL). Within the context of higher education, the relationship between the articulation of ideas in English and student learning outcomes, including academic participation, professional preparedness, and employability, has been widely discussed in previous literature (Richards, 2008). Accordingly, the role of English language learning and teaching in the context of higher education lies in the promotion and development of English-speaking skills, especially in the early stages students' learning pathways.

In EFL contexts such as Vietnam, opportunities for English speaking are largely confined to classroom settings, while real-life communicative use outside the classroom remains limited. Classroom-based English use, including teacher–student interaction, pair work, and group discussions, serves as the primary source of speaking practice for many EFL learners in contexts where opportunities for authentic communication remain limited (Richards, 2008; Littlewood, 2007). These classroom practices are important for the teaching

[1] *English Teacher, FPT University, Hanoi*

of English but are limited to controlled contexts and are not necessarily indicative of the willingness or ability of the students to use English outside the classroom.

Empirical studies suggest that EFL students show a high level of divergence between the use of English for classroom oral communication and the use of English outside the classroom (Littlewood, 2007). In addition to the classroom, the scope for the use of English is limited outside the classroom; therefore, the students are likely to use their first language for daily interactions. In this context, this raises an important question regarding the extent to which classroom speaking practices are transferred to real-life English use. This issue is particularly evident among first-year students. In their early stage of higher learning, many of these students will be enrolled in foundation-level English classes, where their proficiency level ranges from low-intermediate (A2-B1) and their communicative confidence is low. In this transitional period, their affective variables, such as anxiety, fear of making errors, and communicative confidence may play a major role in impeding their willingness to communicate in English, especially in real-world contexts (MacIntyre et al., 1998). They may be highly involved in speaking tasks in class but simultaneously avoid using English outside of class.

While prior studies have examined the teaching of speaking in the classroom from the perspective of EFL settings, not many studies have attempted to examine the ways students use spoken English in various learning contexts, particularly by comparing the use of speaking in the classroom and out of the classroom. In addition, not many studies have attempted to consider the students' own perspectives to account for the prevalence of English speaking in some contexts more than others. In fact, researchers have called for more attention to students' use of language in real contexts to better evaluate the effectiveness of communicative language teaching (Benson, 2011).

While previous studies have extensively explored classroom-based speaking practices in EFL settings (e.g., Bygate, 2001; Nation & Newton, 2009; Goh & Burns, 2012), relatively limited attention has been paid to how instructional speaking practices connect with authentic English use beyond the classroom, particularly from students' perspectives in Asian EFL contexts (Lamb & Arisandy, 2020; Nguyen & Habók, 2023). Specifically, there is a need to bridge the gap between instructed classroom-based speaking and the use of English outside the classroom from the point of view of students themselves, especially first-year Vietnamese university students. To bridge this research gap, the present study aims to explore the ways in which first-year university students employ English for speaking purposes in classroom settings as well as in-class and out-of-class situations. The research aims to explore the ways in which English speaking takes place across different learning environments, as well as the factors that affect English-speaking in and out-of class situations. In line with the research aims, this study employed an explanatory sequential mixed-methods research design to explore English-speaking in EFL university settings. In this regard, the research was guided by the following research questions:

1. How do first-year university students report using English for speaking purposes inside the classroom during a preparatory English program?

2. How do first-year university students report using English outside the classroom in real-life communicative contexts?

3. What factors influence first-year university students' reported use of English-speaking practices outside the classroom?

2. Literature review

2.1 English Speaking in EFL University Contexts

English speaking is considered one of the key aspects of communicative competence in the teaching of second or foreign languages (Canale & Swain, 1980; Richards, 2008). In the context of higher education, the importance of speaking is that it allows the learner to take part actively in the interaction in the classroom, in academic discussions, etc. Speaking competence in EFL contexts involves not only linguistic accuracy but also the ability to communicate fluently and appropriately in meaningful contexts (Bygate, 1987; Richards, 2008).

In many Asian EFL contexts, including Vietnam, classroom interaction remains the primary source of English-speaking practice because opportunities for authentic communication outside the classroom are limited (Littlewood, 2007). As a result, university preparatory English programs often place considerable emphasis on classroom-based speaking activities designed to develop students' basic communicative competence (Hyland, 2006; Richards, 2008). Understanding how students engage in English-speaking practices within and beyond these instructional settings is therefore important for evaluating the effectiveness of EFL speaking instruction.

2.2 Classroom and Out-of-Class English Speaking

In EFL contexts, classroom interaction functions as the primary environment for English-speaking practice because opportunities for authentic communication outside the classroom remain limited (Littlewood, 2007; Benson, 2011). Classroom speaking activities, including teacher–student interaction, pair work, group discussions, and task-based interaction, provide important opportunities for learners to practice and develop spoken English skills (Richards, 2008). However, such interaction is often pedagogically structured and highly task-oriented, with learners responding to teacher prompts rather than engaging in spontaneous communication (Ellis, 2003; Littlewood, 2007). Although these classroom activities may support the development of linguistic competence, they do not necessarily facilitate the transfer of English-speaking practices to real-life communication. This suggests that classroom participation alone may not automatically lead to authentic communicative engagement in EFL contexts due to differing communicative purposes and environmental constraints. Previous studies have suggested that learners who actively participate in classroom-speaking activities may still avoid using English outside the classroom (Littlewood, 2007).

In contrast, out-of-class English speaking typically occurs in less formalized contexts, such as communication with friends, service encounters, or online interaction (Benson, 2011). Recent studies have also explored learners' willingness to communicate beyond the classroom through both human-human and technology-mediated interaction. For example, research has shown that interaction with intelligent personal assistants may increase learners' willingness to communicate outside formal classroom settings by reducing communication anxiety and providing accessible opportunities for English use (Tai, 2024). In many EFL settings, however, such opportunities remain relatively scarce. Benson (2011) argues that learners' out-of-class English use depends largely on both the availability of communicative opportunities and learners' willingness to seek them. As a result, English-speaking practices outside the classroom tend to occur far less frequently than classroom-based speaking activities. Taken together, the literature suggests a substantial discrepancy between classroom participation and authentic English use.

2.3 Factors Influencing English-speaking practices beyond the classroom

Recent studies have pointed to a variety of factors that play a significant role in influencing learners' willingness and ability to use English for speaking outside the classroom

(MacIntyre et al., 1998; Benson, 2011). Among these factors, affective factors have been extensively studied. Collectively, previous studies suggest that affective factors such as anxiety, fear of negative evaluation, and low confidence substantially constrain learners' willingness to communicate in EFL contexts (Horwitz, 1986). Recent studies have also shown that positive affective variables, such as foreign language enjoyment and teacher support, significantly influence learners' willingness to communicate in EFL contexts (Yang et al., 2024). These affective constraints might be compounded in out-of-class situations where learners are more vulnerable to negative social evaluation (Horwitz, 1986; MacIntyre et al., 1998).

The Willingness to Communicate (WTC) framework provides an important perspective for understanding why EFL learners may avoid speaking English outside the classroom despite possessing sufficient linguistic knowledge (MacIntyre et al., 1998). For EFL learners, low levels of willingness to communicate result in the reluctance to engage in English speaking, despite the presence of linguistic competence in the target language (MacIntyre et al., 1998). Similar findings have also been reported in Asian EFL contexts, where learners' communicative confidence and readiness to engage in English interaction is closely associated with confidence and perceived communicative competence (Yashima, 2002). In the Vietnamese EFL context, Nguyen and Vu (2024) similarly reported that learners' willingness to communicate was negatively influenced by affective barriers such as communication anxiety, fear of negative evaluation, and low self-confidence. Besides the role played by affective variables, the impact of contextual or environmental factors cannot be overlooked. For instance, the scarcity of English-speaking interactional partners, the preference to speak the first language, and the absence of communicative need all contribute to the reluctance to engage in real-life interaction in the target language (Littlewood, 2007; Benson, 2011). These factors, in aggregate, result in the reluctance to engage in English speaking outside the classroom (Littlewood, 2007; Benson, 2011).

According to the reviewed literature, the importance of speaking in EFL university education is emphasized, and the classroom is identified as the main setting in which English-speaking practice occurs. The reviewed research has also identified various affective and contextual factors affecting learners' English-speaking practice outside the classroom. However, there are still gaps in the research in this area.

Firstly, despite extensive research on classroom-based speaking instruction in EFL contexts (Richards, 2008; Littlewood, 2007), relatively limited attention has been paid to the discrepancy between classroom participation and real-life English-speaking practices, particularly among first-year university students in Vietnam. Secondly, relatively limited research has focused specifically on first-year university learners enrolled in preparatory English programs, whose communicative confidence and speaking practices may differ from those of more advanced learners due to affective factors such as anxiety and willingness to communicate (MacIntyre et al., 1998). Thirdly, previous studies in this area have often relied primarily on either quantitative or qualitative approaches, while fewer studies have adopted mixed-methods designs to examine English-speaking practices across classroom and out-of-class contexts (Benson, 2011; Littlewood, 2007). In order to fill the gaps in the research, the study employed a mixed-methods research approach to explore the English-speaking practice of first-year university learners in the classroom and outside the classroom, as well as the factors affecting learners' English-speaking practice outside the classroom. The study aims to contribute to the theoretical discussion on how English-speaking practice in the classroom can be promoted to learners to be used outside the classroom.

3. Methodology

3.1 Research design

This study followed an explanatory sequential mixed-methods design, in which quantitative data were collected and analyzed first, followed by the collection and analysis of qualitative data (Creswell & Plano Clark, 2018). The use of a mixed-methods approach was considered appropriate because it enabled the researcher to identify patterns of English-speaking practices across classroom and real-life communicative contexts while also exploring the underlying factors influencing students' English use from the learners' perspectives.

More specifically, the quantitative phase focused on identifying patterns of English-speaking practices inside and outside the classroom, whereas the qualitative phase explored the reasons underlying students' use or non-use of English in real-life communicative situations. The integration of quantitative and qualitative findings provided a more comprehensive understanding of English-speaking practices among first-year university students.

To ensure theoretical grounding, the study was informed by theories of communicative language use and second language communication behavior, particularly Communicative Competence Theory (Canale & Swain, 1980) and the Willingness to Communicate framework (MacIntyre et al., 1998). These theoretical perspectives informed the conceptualization of English-speaking practices as well as the interpretation of the findings.

3.2 Theoretical Framework

The research has been informed by Communicative Competence Theory, as well as the Willingness to Communicate (WTC) model, which provide a unifying framework to guide the research on the phenomenon of English-speaking practices from the perspective of English as a Foreign Language (EFL) context. From the perspective of communicative competence theory, speaking has been conceptualized as "the ability to convey linguistically accurate, fluent, and appropriate language in meaningful contexts, especially in the classroom" (Canale & Swain, 1980; Richards, 2008).

To expand the scope of the study to the wider phenomenon of English-speaking practices beyond the classroom, the WTC model is applied, which posits that language use is shaped by situational factors such as context and the people involved, as well as individual factors such as confidence or anxiety, in the language user (MacIntyre et al., 1998). The synthesis of the two theoretical frameworks provides the potential to offer an explanation of the reason why, despite the linguistic competence, the learner does not engage in real-life communicative situations. These theoretical perspectives informed the interpretation of students' reported speaking practices and the affective and contextual factors associated with English use beyond the classroom.

3.3 Research context and participants

The research was conducted in a private university in Vietnam. English is considered a foreign language in Vietnam. The first-year undergraduate students in the university are expected to take a course in English before they are allowed to specialize in the actual field of study. The course is expected to help the students learn the basics in English, with a focus on speaking. The total number of students in the research was 50. The students in the research were all enrolled in the English course at the time the research was conducted. The level of the students' English was at the A2-B1 level on the Common European Framework of Reference for Language Levels. The research was on a volunteer basis, and all the students were made aware of the confidentiality of the research. The research participants were chosen using a

method of convenience sampling since they were accessible to the researcher during the preparatory English program. Even though such a method of sampling makes the results non-generalizable, it is considered appropriate for exploratory research.

3.4 Data collection instruments

3.4.1 Quantitative data collection methods, in the form of a structured questionnaire, were used to measure the extent of students' use of English in their classrooms and out-of-class activities. The questionnaire had two parts, A and B. In part A of the questionnaire, the background information of the participants, such as their academic background and their level of proficiency in English, was collected. In part B of the questionnaire, the extent of students' use of English in their classrooms was measured by asking them to report their use of English in situations such as answering teachers' questions, participating in pair/group discussions, their use of English in real-life situations outside the school context, etc.

The extent of students' use of English, as reported by the participants of the study, in part B of the questionnaire, is measured by a five- point Likert scale ranging from never to very often. Items in the questionnaire are supported by literature regarding EFL speaking and willingness to communicate. The questionnaire was based on prior research on EFL speaking behaviors and willingness to communicate, as reported by researchers such as MacIntyre et al. (1998) and Benson (2011). To establish content validity, the questionnaire was validated by two EFL lecturers with considerable experience. The pilot study was carried out with a small group of students, consisting of 10 students, to test its reliability and make minor modifications if necessary. The internal consistency of the questionnaire was satisfactory, with a Cronbach's alpha coefficient of 0.82.

3.4.2 Semi-structured interviews

In addition to the quantitative questionnaire, semi-structured interviews were conducted to explore students' perceptions of English use outside the classroom in greater depth. A total of 10 volunteer participants participated in follow-up interviews after the quantitative phase. The interviews focused on students' experiences, attitudes, and perceived difficulties related to using English in real-life communicative situations.

The interviews lasted approximately 15–20 minutes and were conducted in Vietnamese to allow participants to express their perspectives more comfortably and accurately. With participants' consent, the interviews were audio-recorded and later transcribed for thematic analysis. The use of follow-up interviews in explanatory sequential mixed-methods research helped provide deeper explanations for the quantitative findings (Creswell & Plano Clark, 2018). In addition, the number of interview participants was considered appropriate for exploratory qualitative inquiry and thematic analysis, where relatively small samples may still provide sufficient thematic depth and saturation (Braun & Clarke, 2006; Guest et al., 2006).

3.5 Data collection procedure

The data collection was done in the middle of the English preparatory program. This ensured a high rate of return because it was administered in class. It took the participants about 20 minutes to answer the questionnaire. After analyzing the quantitative data, the semi-structured interviews were conducted. This ensured that the qualitative part explained the quantitative part.

3.6 Data analysis

Descriptive statistics, including frequencies, percentages, means, and standard deviations, were used to analyze quantitative data and identify patterns of English-speaking practices across contexts. In addition, a Wilcoxon signed-rank test was conducted to compare students' reported English-speaking frequency inside and outside the classroom, as the data involved paired ordinal Likert-scale responses from the same participants. In terms of qualitative data analysis, thematic analysis was used following the six-phase procedure suggested by Braun and Clarke (2006), which involved familiarization with the data, generating codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. This helped to identify major patterns related to affective and contextual issues in the students' use of English outside the classroom setting. Data collection and analysis continued until thematic saturation was reached, that is, when no substantially new themes or patterns emerged from the interview data (Guest et al., 2006). Recurrent ideas became increasingly repetitive across participants, suggesting sufficient thematic coverage for the purposes of the study. To enhance trustworthiness, the researchers discussed and refined the coding categories collaboratively throughout the analysis process.

3.7 Ethical Considerations

Ethical consideration was strictly adhered to in the study. Participants were made aware of the study's objectives, as well as their right to withdraw at any time with no penalty. Informed consent was obtained before data collection, with confidentiality of the collected data being ensured by anonymizing the information collected. Data collection was limited to research purposes only. All participants provided informed consent prior to participation, and no personal identifying information was collected to ensure anonymity.

4. Findings

4.1. Questionnaire findings

4.1.1 English-Speaking Use inside the Classroom

The results of the questionnaire indicate that English speaking was frequently used inside the classroom during the English preparatory program. Overall, students reported high levels of engagement in English-speaking activities inside the classroom, especially when these activities were teacher-initiated.

A majority of participants (78%) reported frequently using English to answer teachers' questions. The activity was found to be used most frequently compared to all other speaking activities, with a mean score of M = 3.96 (SD = 0.61) on a five-point Likert scale. In parallel, it was found that 72% of students (n = 36) reported that they frequently took part in pair or group discussions held in English.

In contrast, it was found that speaking activities that required higher spontaneity were used less frequently. Only 34% of participants (n = 17) reported that they frequently volunteered to give their opinions without being prompted by teachers, whereas 46% (n = 23) reported that they seldom did so. The mean score was M = 2.91 (SD = 0.83).

These findings reveal that English speaking in the classroom was primarily facilitated by the teacher and was task-based, as students relied on instructional cues to engage in English speaking. Although English speaking was highly facilitated by opportunities provided by the classroom, these opportunities were primarily task-based and embedded in instructional contexts rather than students' spontaneous English speaking.

Classroom English-speaking activities	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Very often (%)	Mean	SD
Responding to teachers' questions	2 (4%)	4 (8%)	5 (10%)	21 (42%)	18 (36%)	3.96	0.61
Participating in pair/group discussions	3 (6%)	6 (12%)	5 (10%)	20 (40%)	16 (32%)	3.74	0.67
Giving prepared oral responses	4 (8%)	7 (14%)	10 (20%)	18 (36%)	11 (22%)	3.50	0.79
Volunteering opinions without prompts	8 (16%)	15 (30%)	10 (20%)	12 (24%)	5 (10%)	2.91	0.83

Table 1

Frequency of English-Speaking Use inside the Classroom (n = 50)

As illustrated in Table 1, English speaking inside the classroom was most often engaged in during teacher-initiated activities. Responding to teacher-posed questions was the most common form of English speaking, achieving the highest mean score (M = 3.96, SD = 0.61), while participating in pair or group discussions came second (M = 3.74, SD = 0.67). English speaking engaged in for the purpose of self-initiating, such as volunteering opinions without any prompts, was reported the least (M = 2.91, SD = 0.83).

4.1.2 English-Speaking Use outside class in Real-Life communicative Contexts

In contrast to the school context, the use of English outside the school was found to be limited. The data collected by the questionnaire show that there was a considerable decrease in the rate of English use when the contexts of real-life English-speaking activities were considered.

In the outside class context, only 18% of the participants (n = 9) reported that they used English frequently or very frequently, whereas 62% of the participants (n = 31) reported that they used English rarely or not at all for speaking activities in daily life. The overall mean score for the use of English outside the school for speaking activities was M = 2.14 (SD = 0.71), which was significantly lower than the score for English-speaking activities in the school context.

However, a closer examination of the data indicated that 70% of the students (n = 35) used English infrequently when interacting with their peers, including those who were simultaneously taking English courses. Additionally, 76% of the students (n = 38) stated that they used English almost never when interacting in service encounters, such as ordering food, shopping, or seeking information. The use of English in online communication was slightly higher but still limited, with 42% (n = 21) of the students stating that they used English occasionally.

The above data suggest that there was a significant gap in the use of English in the classroom and in out-of-class contexts. Although the students used English frequently in the

classroom, they used English minimally in real-life communicative contexts. A Wilcoxon signed-rank test revealed a statistically significant difference between students' reported English-speaking frequency inside and outside the classroom, with classroom speaking reported more frequently, $Z = -5.42$, $p < .001$, $r = .77$, indicating a large effect size.

Out-of-class English-speaking contexts	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Very often (%)	Mean	SD
Communicating with friends in English	15 (30%)	20 (40%)	8 (16%)	5 (10%)	2 (4%)	2.18	0.88
Using English in service encounters (e.g., shopping, ordering food)	18 (36%)	20 (40%)	7 (14%)	4 (8%)	1 (2%)	2.00	0.82
Using English in everyday interactions outside class	17 (34%)	21 (42%)	8 (16%)	3 (6%)	1 (2%)	2.00	0.80
Using English in online communication	10 (20%)	14 (28%)	11 (22%)	10 (20%)	5 (10%)	2.72	1.12

Table 2

Frequency of English-Speaking Use outside class (n = 50)

As shown in Table 2, English use outside the school environment was found to be generally low in most real-life situations. The lowest mean scores were found in service encounters and daily interactions, which had a mean of 2.00 each, implying that most students made little to no use of English in these situations. Although English use in online communication seemed to be a bit more frequent, with a mean of 2.72 and a standard deviation of 1.12, it is still low compared to English use in speaking activities in the classroom. In conclusion, there is a big gap between English-speaking use in the classroom and English use outside the school environment.

4.2 Qualitative Findings from semi-structured interviews

A total of 10 out of the 50 participated in the semi-structured interviews. The answers were analyzed using thematic analysis to investigate the perceptions of the use of English outside the classroom and the quantitative patterns. Two major themes were identified from the answers: affective factors and contextual/environmental factors.

4.2.1 Affective Factors Influencing English-Speaking Use

The affective factors proved to be the most important factor in explaining the students' minimal usage of English outside of academic settings. Out of the ten students, seven students pointed out that their lack of confidence in speaking English was the major factor in not using English in everyday life. Some of the students had concerns about their English pronunciation, vocabulary, and grammar, which made them reluctant to speak English unless they had to. One student said: *"I am not confident about my pronunciation, so I usually avoid speaking English outside the classroom."* (Student 3)

Another student similarly noted:

"I worry that I might use wrong words or grammar, so I feel safer not speaking English." (Student 7)

Moreover, six of the ten students cited fear of errors or negative evaluation as a significant factor in their response to barriers. These students saw the level of stress in communication in the real world as greater than in the classroom, where making errors is acceptable. One of the students explained:

"In class, making mistakes is normal, but outside the class I am afraid people will judge my English." (Student 1)

Another student highlighted the anxiety associated with speaking to unfamiliar interlocutors:

"When I speak English with strangers, I feel very nervous because I don't want them to think my English is bad." (Student 9)

These responses indicate that affective constraints—particularly low confidence and fear of negative evaluation—played a crucial role in shaping students' English-speaking behavior outside school, despite their active participation in classroom speaking activities.

4.2.2 Contextual and Environmental Factors

In addition to affective factors, students also identified contextual factors that constrained their opportunity to use English outside of class. Of the 10 students, eight students identified that they had limited opportunity to use English in daily life because the majority of interactions were conducted in their native language. Some of the students emphasized that English was not required outside of class, which affected their motivation to initiate English-speaking interactions. One student commented:

"In daily life, everyone speaks Vietnamese, so there are not many chances to use English." (Student 4)

Another student explained:

"Outside school, English is not really necessary, so I don't think about using it." (Student 10)

Moreover, five of the ten students identified the lack of English-speaking interlocutors as an important barrier. Although the students were willing to practice their English skills, they encountered difficulties in locating interlocutors who were willing to communicate in English.One participant noted:

"I want to practice English, but my friends prefer speaking Vietnamese." (Student 6)

Another student echoed this concern:

"It is hard to find someone to talk to in English, so I usually use my first language." (Student 8)

Together, these contextual and environmental factors reinforced the gap identified in the quantitative findings, constraining students' English-speaking use beyond the classroom.

5. Discussion

This study is concerned with exploring the students' English-speaking practices in the classroom and in real-life situations in an EFL context, as well as the factors that affect students' engagement in English speaking activities outside the classroom. The findings show that there is a large gap between students' English speaking in the classroom and in real-life situations, which can be explained in terms of communicative competence and WTC.

5.1 Classroom English Speaking as Instructed Communicative Practice

Moreover, the findings of the questionnaire show that English-speaking frequently occurred in classroom settings, particularly in activities initiated by teachers, such as answering teachers' questions and participating in pair or group discussions. This finding suggests that classroom speaking remains largely pedagogically controlled, limiting the development of spontaneous communicative competence, which aligns with Littlewood's (2007) distinction between form-focused and meaning-focused interaction. The slightly higher mean scores in these activities suggest that the English preparatory program provides students with opportunities to practice English-speaking in the classroom.

The low incidence of self-initiated oral contributions, such as volunteered opinions may imply that English speaking in the classroom was still largely controlled. From the point of view of communicative competence (Canale & Swain, 1980; Richards, 2008) these classroom practices may promote the development of linguistic accuracy and fluency but may not necessarily facilitate spontaneous communicative use in authentic contexts. Thus, the findings imply that while students were engaged in English speaking activities, these activities were still largely in the context of instruction, not communicative need.

5.2 Limited Transfer of Classroom Speaking to Real-Life communicative contexts

A key finding of the study is the markedly low frequency of English-speaking use outside school. The majority of students reported very little or no use of English in out-of-class situations, such as interactions with peers or service encounters. This finding supports the claims made in earlier research that classroom ability does not necessarily equate to the use of the target language in other situations (Littlewood, 2007).

The slight increase in the use of English in online communication may be an indicator of the potential of these online contexts to be relatively safe avenues for English-speaking activities. However, the overall use of out-of-class English language use is still restricted, which aligns with the notion that English language use is often perceived by EFL learners as an academic subject. This aligns with Benson's (2011) notion that out-of-class language use in the context of EFL settings is heavily dependent on the opportunity and the agency of the learners.

5.3 Affective Constraints and Willingness to Communicate

Qualitative findings provide substantial insights into the factors that prevent students from using English outside of the classroom. Among the ten students, low self-confidence and fear of making mistakes dominated the affective factors that prevent students from using English outside of the classroom, which is consistent with the literature of FLA that indicates that fear of negative evaluation is one of the most important factors that prevent students from using English, according to Horwitz (1986). Similar patterns have been identified in Asian EFL classrooms, where students' willingness to communicate is strongly influenced by confidence, classroom climate, and fear of negative evaluation (Peng & Woodrow, 2010).

Interpreted in the light of the WTC framework proposed by MacIntyre et al. (1998), the results suggest that the willingness to use English outside the classroom is not limited by the lack of linguistic knowledge, but rather by other factors. For instance, the students who were actively using English in the classroom were still reluctant to use it outside the classroom due to the risk of communication. This is in support of the contention that linguistic knowledge is not enough to account for the use of the target language, especially in EFL settings.

5.4 Contextual Constraints and Lack of Communicative Necessity

Apart from affective factors, contextual and environmental factors also limited students' English-speaking practices. The major part of interviewed students emphasized that most interactions were based on the first language and that there was no need to use English outside the school environment. This result corresponds to previous research that showed that lack of exposure and absence of communicative need reduce students' motivation to use the target language outside the classroom (Littlewood, 2007; Benson, 2011).

The difficulty in finding an English-speaking interlocutor again illustrates the ways in which the environmental factors work in tandem with the affective factors to create an unwillingness to communicate in English. In the WTC framework, the role of the contextual factors is to reduce the learner's situational willingness to communicate, despite the learner's underlying motivation to improve his/her speaking skills in English. The interplay between the affective and contextual factors explains the continued gap between the learner's use of English in the classroom and in the real world, as seen in the results of this study.

5.5 Pedagogical Implications

The results of this research also reveal several implications for English language teaching pedagogy in preparatory English programs offered in EFL university contexts. Firstly, while speaking activities in the classroom are undoubtedly important, more attention needs to be devoted to overcoming affective factors by creating a speaking environment that encourages risk-taking and self-initiated speaking. Secondly, it may be necessary to take English language teaching beyond the confines of the classroom by designing speaking tasks that involve out-of-class speaking.

Finally, promoting students' understanding that English is not just an academic subject but also a viable means of communication may enhance their interest in communicating beyond the classroom. Such changes in teaching methods may help bridge the gap between instructed language use and communicative language use.

5.6 Summary

In summary, the findings highlight a substantial limited correspondence between instructional speaking activities and spontaneous English interaction which is shaped by the interaction of instructional, affective, and contextual factors. While the classroom provides learners with the opportunity to practice speaking in English frequently, it is important to note that the opportunity to practice speaking in the classroom does not necessarily equate to the real-life setting.

5.7. Limitations

In spite of the positive contributions of the study, several limitations need to be pointed out.

Firstly, this study had a small sample size of 50 participants from a private university in Vietnam. This may limit the generalizability of the findings to broader EFL contexts. Secondly, the present study is based on self-report measures of the participants. The responses gathered by the survey technique are prone to the influences of recall and social desirability biases. Therefore, one cannot be sure that they reflect fully the students' actual behavior of using English to communicate. Future research should triangulate self-reported questionnaire and interview data with more objective sources of evidence, such as classroom observations, speaking performance assessments, audio/video recordings of speaking tasks, and teacher evaluations.

Thirdly, although the study had a mixed-methods design, the qualitative part of the study had a small sample size of 10 participants. This might not have been representative of the entire range of the participants' experiences. In addition, the study focused more on the frequency of the participants' English-speaking behavior rather than the effectiveness of the communication. This is another limitation of the study.

Lastly, this study is a snapshot of the participants' behavior and cannot account for changes in the participants' behavior over time. Future studies might employ longitudinal designs to assess the development of the participants' English-speaking behavior. Despite these limitations, the study provides important insights into the gap between classroom-based and real-life English-speaking practices in EFL contexts.

6. Conclusion

This study examined the speaking practices of first-year university students both inside the classroom and in real-life situations in the context of an English as a Foreign Language (EFL) classroom. The study's findings reveal a wide gap in the level of participation in speaking practices inside the classroom and the limited use of such practices outside the classroom. Despite the opportunities for speaking practice inside the classroom, such practice may not easily extend to authentic communicative contexts. The study's findings also reveal the importance of affective factors in the use of English outside the classroom. The low level of confidence, the fear of making mistakes, and the anxiety stemming from negative evaluation all served to limit the students' willingness to communicate. The lack of opportunities and the lack of communicative necessity also limited the students' use of English outside the classroom.

From the perspective of language teaching, the study's findings reveal the importance of moving beyond the classroom-bound speaking practice to more authentic, meaningful, and confidence-building speaking practice. The study's findings also reveal the importance of language teaching going beyond the development of students' linguistic knowledge to the development of their willingness and readiness to communicate in real-life situations.

The study's findings also have important implications for the development of students' readiness to communicate in the global academic or professional community, where the ability to communicate orally plays a crucial role. Future studies on the speaking practices of students need to be longitudinal studies to reveal the development of students' speaking practices over time.

REFERENCES

Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). Routledge.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bygate, M. (1987). *Speaking*. Oxford University Press.

Bygate, M. (2001). Speaking. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 14–20). Cambridge University Press.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/1.1.1

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. Cambridge University Press.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly, 20*(3), 559–562. https://doi.org/10.2307/3586302

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

Lamb, M., & Arisandy, F. E. (2020). The impact of online use of English on motivation to learn. *Computer Assisted Language Learning, 33*(1–2), 85–108. https://doi.org/10.1080/09588221.2018.1545670

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching, 40*(3), 243–249. https://doi.org/10.1017/S0261444807004363

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a second language: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.

Nguyen, T. T. H., & Habók, A. (2023). Vietnamese EFL students' willingness to communicate in English and its related factors. *Heliyon, 9*(5), e15834. https://doi.org/10.1016/j.heliyon.2023.e15834

Nguyen, V. H., & Vu, H. N. (2024). A phenomenological analysis of the obstacles impeding Vietnamese EFL learners' willingness to communicate in English. *International Journal of English Language Studies, 6*(2). https://doi.org/10.32996/ijels.2024.6.2.21

Peng, J.-E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning, 60*(4), 834–876. https://doi.org/10.1111/j.1467-9922.2010.00576.x

Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.

Tai, T. Y. (2024). *Comparing the effects of intelligent personal assistant-human and human-human interactions on EFL learners' willingness to communicate beyond the classroom*. *Computers & Education, 210*, 104965. https://doi.org/10.1016/j.compedu.2023.104965

Yang, Y., Cui, Y., & Yao, S. (2024). Teacher support, grit and L2 willingness to communicate: The mediating effect of foreign language enjoyment. *BMC Psychology, 12*, 383. https://doi.org/10.1186/s40359-024-01877-5

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86*(1), 54–66. https://doi.org/10.1111/1540-4781.00136

Author's bio:

Nguyen Thi Ngoc Chinh is an English lecturer at FPT University, Vietnam, with 15 years of experience teaching university students. Her teaching and research interests include EFL speaking, critical reading, and communicative language teaching in higher education contexts.

#18 Exploring University Students' Use of AI Tools in Academic Writing at HUFLIT

Phan Ngoc Phuong Vy[1]

Abstract

The past few years have witnessed the rapid emergence of generative artificial intelligence (AI) tools which have significantly reshaped the academic writing processes in higher education. There has been an increasing tendency among the undergraduates to utilise AI-powered platforms to support multiple stages of academic writing, including brainstorming, initial drafting and revision processes. This study is aimed to explore how university students describe their experiences of employing AI-driven tools in their academic writing works and their attitudes towards AI's position in reconstructing their writing practices. Adopting a qualitative research design, this study relied on semi-structured interviews conducted with first-year students at HUFLIT who frequently integrate AI tools in their academic writing tasks. The collected data underwent thematic analysis to identify recurring patterns that reflect their experiences and interpretation for their engagement with AI in writing skills. The findings reveal that the participants perceived AI as a form of cognitive scaffolding which supported idea generation, language refinement and structural organisation throughout their academic writing tasks. Concurrently, the students expressed concerns about the over-reliance on AI tools, the difficulties in identifying authorship and the academic integrity. Lastly, the results suggest that students had mixed views of AI's impacts on their future writing development while attempting to balance between efficiency brought by AI and their true skill development. This study expects to contribute profound understanding of responsible AI integration and relevant pedagogical approaches to academic writing skills in the era of generative AI.

***Keywords*:** generative AI, academic writing, EFL, university students, qualitative study

1. Introduction

The widespread expansion of generative AI tools, such as ChatGPT, has substantially transformed academic writing practices in higher education, influencing the way learners brainstorm ideas, organise texts and refine their language (Zhao et al., 2024). Recent studies on generative AI highlight its increasing integration into student academic wiring class, suggesting that these technologies are not only used for surface-level editing, but also function as cognitive assistance in more complex writing processes (Sanz-Tejeda et al., 2026). Besides, syntheses of research in AI-integrated writing suggest that students employed these tools to enhance coherence, lexical quality and textual organisation; at the same time, these researchers also identified a number of potential drawbacks, including growing dependence, ethical concerns and deterioration in metacognitive development (Chanpradit, 2025; Sanz-Tejeda et al., 2026).

In addition to the reported efficiency in improving writing performance, students also hold varied views towards AI tools. Particularly, qualitative studies exploring students' perspectives on AI tools indicate that AI is expected to function diverse pedagogical roles, from a drafting assistant to a virtual tutor facilitating different phases of the writing process (Kim et al., 2024). Such findings strengthen AI's potential in supporting students' writing activities

[1] *MA, Faculty of Foreign Languages (FFL), HUFLIT, Ho Chi Minh City, Vietnam; Email: vypnp@huflit.edu.vn*

while also revealing its challenges in applying AI academic tasks in general. Although a growing literature of research has explored students' patterns of AI usage and their general perceptions of these technologies, many of these current studies are largely constructed on survey-based and quantitative methodologies, thereby paying little attention to the depth and complexity of students' lived experiences with AI integration in their academic writing phases. Besides, this limitation in methodologies also resulted in limited insight into students' interpretation of AI's role in shaping their academic writing skills, particularly in Vietnam where national policies have increasingly emphasised the importance of transforming English education from a traditionally foreign-language system into a second-language oriented framework.

Furthermore, most existing studies are situated within general higher education or non-EFL environments, resulting in a significant research gap in understanding about how learners in EFL university context use AI and interpret its role in academic writing. To address this gap, this study adopts a qualitative interview design with university students at the Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT) to answer two research questions:

1. How do university students at HUFLIT describe their experiences of using AI tools in academic writing?

2. How do university students at HUFLIT make sense of the role of AI tools in shaping their academic writing practices?

By focusing on students' perspectives in a Vietnamese EFL university context, this research aims to deepen understanding of AI-mediated academic writing activities and offer pedagogical approaches that balance technological benefits and ethical considerations in EFL higher education.

2. Literature Review

2.1. AI as Cognitive and Linguistic Scaffolding in Academic Writing

The increasing growth of generative AI tools has substantially transformed academic writing practice within higher education institutions. In particular, recent studies argue that AI-based platforms, such as ChaGPT, operate not merely as editing tools, but as cognitive assistance that facilitates idea generation, organisation, drafting and iterative revision (Kasneci et al., 2023; Dwivedi et al., 2023; Zhai, 2023). Unlike earlier grammar-check softwares, generative AI tools are capable of generating extended and coherent responses, which allows learners to interactively engage in the drafting and language refinement processes (Rudolph et al., 2023). As a result, these tools are widely recognised as forms of technological scaffolding that mediate cognitive processes during complex writing tasks (Farrokhnia et al., 2023).

In fact, the writing process includes several repeated phases, such as idea planning, composing and draft reviewing (Flower & Hayes, 1981). Generative AI tools can be employed to facilitate these stages by providing prompts, suggesting structures and improving sentence formulation (Yan, 2023; Wang et al., 2024). Such support is consistent with earlier studies on the digital writing environment, which suggest that technology can influence not only the textual output produced by writers but also the cognitive strategies writers employ when composing (Warschauer, 2010). Therefore, AI-assisted writing can be seen as a higher stage of computer-mediated composition, where technology supports both language correction and complex idea development.

The scaffolding function above may be especially important in EFL contexts, where the act of writing in a foreign language demands greater cognitive load due to restricted lexical resources, grammatical uncertainty and unfamiliar disciplinary norms (Hyland, 2019). Within such settings, AI-generated feedback can reduce linguistic burden and support the development of fluency in writing (Alhusaiyan, 2024; Jacob et al., 2024). Moreover, a recent empirical study show that students perceive AI as an integrated writing partner, rather than a tool offering only surface support such as idea generation, organization and iterative revision (Kim et al., 2024; Woo et al., 2026). While Kasneci et al. (2023) emphasise the pedagogical benefit of large language models in fostering learner autonomy, Dwivedi et al. (2023) indicate that the effective impact of AI on the learning process depends on how students integrate these tools into their writing practices. In summary, these studies suggest that AI has the capacity to improve both academic writing performance and cognitive development throughout the whole writing process itself. However, to capture the nature of this transformation, it is a must to explore how students describe their experiences with AI and interpret its role in enhancing writing competence.

2.2. Students' Conflicting Feelings about Using AI in Writing

Despite the widespread adoption of these AI tools, students' perceptions of their roles are often complex and sometimes ambivalent. In particular, a number of empirical studies report that many learners view generative AI as beneficial tools that improve clarity, lexical diversity and structural coherence (Baidoo-Anu & Ansah, 2023; Chan & Hu, 2023). Specifically, learners commonly explain they use AI to overcome writer's block, outline structures and paraphrase ideas, especially in ESL or EFL contexts. From this perspective, AI is frequently conceptualised as a supportive assistant that enhances writing efficiency and learner confidence.

Simultaneously, current research has raised concerns regarding over-reliance, academic integrity and authenticity of ownership (Cotton et al., 2023; Baidoo-Anu & Ansah, 2023). In particular, Cotton et al. (2023) argue that the extended accessibility of generative AI tools disrupts the traditional understanding of originality, and thereby creating challenges for educational organisations to maintain academic integrity. Consistent with the aforementioned studies, Chan and Hu (2023) also report that students appreciate the benefits of AI yet remain uncertain about its ethical limits. Similarly, within EFL settings, students indicate that ChatGPT – a common AI-based tool – contributes to improved writing performance, reduced anxiety and greater confidence; however, they also express concerns about excessive dependence and ownership (Teng, 2024; Barzanji & Alsofyani, 2026). This implies that students do not use AI in a simple and practical way. In other words, students question whether AI supports their learning or threatens the originality of their writing. Such complicated perceptions highlight the need to understand how learners interpret the adoption of these mediating technologies in their writing practices, rather than simply focusing merely on how frequently they use these tools.

2.3. Shifting the Focus from Outcomes to Student Perspectives

Although the interest in AI-assisted writing practice is increasingly growing, most previous studies concentrate on student performance, institutional policies and university reactions, rather than students' personal experience. A number of investigations address the potential transformation that educational systems may undergo and suggest future strategies in the era of generative AI (Dwivedi et al., 2023; Kasneci et al., 2023). In the meantime, others concentrate on academic integrity and regulation (Cotton et al., 2023). While these

contributions are valuable to understand policies and institutional settings, they pay less attention to how learners actually experience and integrate AI in their writing practices.

A number of survey-based research seem useful when identifying students' general attitudes towards AI-mediated writing processes (Baidoo-Anu & Ansah, 2023; Chan & Hu, 2023). However, these quantitative-designed studies often view these attitudes as numbers and fixed variables, rather than as emotions and experiences that may change and develop over time. In fact, academic writing activities which are socially shaped allow students to develop and express their understanding. As a result, if AI is used for initial drafting and revision, it may affect how students perceive authorship and their responsibilities. To fully understand this shift, qualitative research is necessary to explore students' detailed reflections on their experiences with AI.

2.4. AI Use in EFL Higher Education and the Need for Context-Sensitive Inquiry

The need for qualitative research becomes even stronger in EFL higher educational contexts. Academic writing in the EFL context is actually challenging due to linguistic limitations and academic standards (Hyland, 2019). In such cases, AI may function both as a tool to increase productivity and a linguistic support that influences their confidence and independence as writers. Although Kasneci et al. (2023) emphasise the potential of AI to support multilingual learners, and Dwivedi et al. (2023) highlight the importance of context-awareness when engaging AI in education, qualitative research examining students' experiences on AI-assisted writing in the EFL university settings remains limited.

2.5. Research Gap

In summary, existing studies suggest that generative AI-based tools in academic writing can serve as a form of cognitive support (Kasneci et al., 2023; Dwivedi et al., 2023) that students hold mixed views towards its use (Baidoo-Anu & Ansah, 2023; Chan & Hu, 2023); concurrently, these research shows that concerns about ethics and institutional policies are increasing (Cotton et al., 2023). Despite this fact, qualitative research exploring university EFL students' experiences and interpretations of AI's role in their academic writing mastery remains limited. Therefore, to address this gap, this study aims to investigate how students describe their engagement with AI in academic writing tasks and how they understand its impact on their academic writing development.

2.6. Theoretical Framework

This study is grounded in sociocultural theory (SCT), which views learning as a process that develops through social interactions with cultural tools within a particular environment (Vygotsky, 1980). From this viewpoint, cognitive development arises through interactions with mediational resources such as language, technologies and social relationships, rather than occurring solely within the individual's mind. In educational settings, such supports play a significant role in shaping the way learners engage with tasks, build knowledge and develop new competencies.

A key concept in SCT is mediation referring to the way how human cognition is shaped and supported by symbolic tools and cultural artifacts (Thorne & Lantolf, 2006). In second language acquisition, mediation commonly occurs through interactions with teachers, peers and instructional materials that guide and influence learners' performance. As digital technologies are increasingly present in higher education, they can function as mediational tools that shape the learning processes. Therefore, in the mastery of academic writing skills,

generative AI tools may serve as an effective support in generating ideas, organising arguments and refining language use.

Closely linked to mediation is the concept of the Zone of Proximal Development (ZPD), which mentions the gap between what learners can independently achieve and what they can accomplish with guidance and assistance (Vygotsky, 1980). Traditionally, such support within the ZPD has been offered by more knowledgeable individuals; nevertheless, these modern digital technologies are able to provide forms of assistance that is similar to guided support. As a result, suggestions, feedback and prompts created by AI-based tools may enable students to complete writing tasks that exceed their current linguistic or cognitive abilities.

Based on Vygotsky's SCT, another key concept is scaffolding built by Wood, Bruner, and Ross (1976), which describes temporary support provided to learners when they are engaged in tasks beyond their current capacities. In particular, within the setting of academic writing, AI tools may serve as scaffolding by helping students organise ideas, improve sentence formation and revise drafts. However, SCT highlights that tools do not automatically determine learning outcomes. Instead, their influence depends on how learners adopt and integrate them into their learning practices (Thorne & Lantolf, 2006). Guided by this perspective, this study explores how university students at HUFLIT describe their experience with AI tools and how these technological tools shape their academic writing practices.

3. Methodology

3.1. Research Design

This study employs a qualitative design to investigate university students' experiences with AI-powered tools in academic writing, because the qualitative approach is particularly appropriate for fully understanding how individuals interpret and make sense of their experiences within a social or educational context (Creswell & Poth, 2018). Semi-structured interviews were adopted to collect data, as this approach allows the researcher to obtain detailed description of participants while enabling flexibility to prompt ideas emerging from the interviews (Kvale & Brinkmann, 2015).

3.2. Participants

The sample consisted of 18 first-year university students majoring in English at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). Participants were recruited using a purposive sample, which merely choose those individuals who have relevant experiences to the phenomenon being studied (Patton, 2014). First-year undergraduate students were selected, because they are in the early stages of developing academic writing competency and are increasingly exposed to AI learning tools. The students were first recruited on a voluntary basis, and those who reported having used AI tools for learning and writing activities before would be invited to participate in semi-structured interviews. The sample size was determined according to the principle of data saturation, which occurs when no new themes or meaningful patterns emerge (Guest et al., 2006). Although the participants were first-year undergraduate students, they had completed several months of university study and academic writing practice prior to the interviews, enabling the study to capture emerging shifts in their use of AI tools and their engagement with AI-assisted writing.

3.3. Data Collection and Analysis

This interview protocol examined participants' motivations for AI usage, the particular phases in which these technologies are utilised, students' interactions with AI-generated

content, the initial transition in their behaviours and their reflections on authorship, responsibility and skill development.

The data were analysed using the thematic analysis approach proposed by Braun and Clarke (2006). The analysis involved several stages, namely familiarisation with the data, generalisation of initial codes, identification of recurring patterns and development of broader themes. This process allowed the author to identify the key themes reflecting student experiences with AI tools and their interpretation of AI tools' roles in their academic writing practices.

3.4. Ethical Considerations

Ethical considerations were carefully addressed throughout the research process. Participation was voluntary, and all students were informed about the purpose of the study before the interviews. All their responses were kept confidential and used only for the research purpose. To ensure anonymity, all participants were addressed as P1, P2, P3… during the analysis and reporting.

3.5. Theme Emergence and Data Saturation

Initial patterns appeared during the first few interviews, in which participants consistently described AI as a helpful tool that supports idea development and language refinement. As the interviews continued, more detailed patterns became evident, especially a shift from early dependence on AI-generated content to a more strategic and selective use. Data saturation was reached between Participants 12 and 15, as no new significant themes were identified after this stage. The subsequent three interviews were conducted to confirm the stability and comprehensiveness of the thematic framework, with patterns related to writing practices, AI use, and ethical issues.

4. Results

The thematic analysis of the interview data from 18 participants revealed four major themes that reflect students' experiences with AI tools in academic writing. Responses showed a high degree of consistency across participants, indicating similar patterns emerged among participants.

4.1. Theme 1: AI as a Multifaceted Writing Assistant

Across the dataset, all participants consistently conceptualised AI as a supportive "assistant" that facilitates linguistic and technical issues of writing practices, rather than substitutes human cognition. These students refer to AI as a multifaceted writing assistant that is effectively engaged in various stages of the writing process.

At the initial stage of writing, all participants reported utilising AI for brainstorming and outlining, especially to handle "writer's block" (P2) and build an initial structure. For example, Participant 2 depicted AI tools as "incredibly helpful for outlining, checking grammar, and refining vocabulary" while Participant 12 explained that they heavily relied on AI because they "found it difficult to create a clear outline and use formal academic language."

In the later stages of the writing process, AI was widely employed to proofread and refine language. In addition, participants reported relying on AI-based tools to enhance grammatical accuracy and strengthen the academic tone of their writing. In short, according to these findings, AI provides ongoing support, and thereby proves to be effective in both idea generation and language refinement.

4.2. Theme 2: Negotiating Dependence and Critical Engagement

As can be inferred from the dataset, students' use of AI showed a clear development over time. While early use of AI was characterised by a level of dependence, their later practices reflected a more strategic and critical engagement of AI in the writing tasks.

In the initial phase, a large number of participants (16 out of 18) admitted using AI to produce extended text, reflecting a lack of confidence in their writing capacities. For instance, Participant 7 noted that "at first, I used AI mainly for generating paragraphs because I lacked confidence in my academic writing."

As students gained more experience in writing skills, their use of AI became increasingly selective and purposeful. Instead of allowing AI to complete the whole writing practices, they now merely used it to evaluate coherence and improve language structure. To illustrate, Participant 2 described this change, noting that AI is used as "a sophisticated research assistant" to help with "logic checking and sentence restructuring". Likewise, Participant 15 highlighted a move towards greater independence, saying that they use "only for ideas and grammar" in order to support their writing development. This shift reflects increased learner agency and metacognitive awareness in AI-assisted writing.

4.3. Theme 3: Psychological Empowerment: Confidence and Reduced Anxiety

Besides cognitive support, AI was linked to positive psychological outcomes. In particular, participants consistently reported that AI increased their confidence and reduced stress when doing academic writing practices, especially in foreign language context.

Several participants (e.g. P1, P2, P4, P8, P12, P14, P15, P18) indicated that their confidence in their writing had enhanced. Particularly, Participant 2 reported that "AI has significantly refined my language and boosted my confidence"; in the meantime, other participants noted that they felt more confident about the accuracy and clarity of their writing.

Furthermore, many students (e.g. P4, P7, P13, P15, P18) also highlighted the role of AI in reducing anxiety. For example, Participant 7 explained that "knowing I can revise and refine my work with support reduces anxiety", making their academic work "more efficient and less stressful". Overall, these findings suggest that AI tools support writing performance and boost emotional experience in academic writing.

4.4. Theme 4: Assertion of Authorship and Ethical Boundaries

Despite their frequent use of AI in the academic writing process, students showed a strong and clear awareness of authorship and academic integrity. What is worth mentioning is that all participants emphasised that they remained responsible for their own work.

Many participants consistently identified themselves as the primary authors who retain control over ideas, arguments and final decisions. For instance, Participant 14 stated "the responsibility for the arguments, structure, and accuracy of the work remains mine" while Participant 2 highlighted that "AI only acts as a tool to polish my original work".

Moreover, many students actively set clear ethical boundaries for their use of AI. Many participants reported avoiding AI in reflective writing tasks to maintain authenticity and personal voice. To illustrate, Participant 7 noted that they "avoid using it for highly personal reflective essays" to ensure authenticity. Similarly, a large percentage of students (except P2 and P6) noted that they refused to use AI during examinations in order to comply with the institutional regulations. In short, these findings suggest student self-regulation and ethical awareness when using AI in their work.

4.5. Summary

Overall, the findings above reveal a consistent pattern in which AI is perceived as a supportive tool that enhances both cognitive and emotional dimensions of academic writing. At the same time, students strongly maintain awareness of authorship and thereby actively regulate their use of AI in line with personal and academic integrity standards.

5. Discussion

The findings of this study offer important insights into how students experience and interpret the use of AI tools in academic writing in higher education context. From the perspective of Sociocultural Theory (SCT), the results can be explained through key concepts, such as mediation, scaffolding and learner agency while also reflecting the features of EFL contexts in Vietnam in particular and Asian countries in general.

5.1. AI as a Mediational Tool and Scaffolding Mechanism

In response to the first research question, the results suggest that students consistently view AI as a multifaceted writing assistant that facilitates various stages of the writing process, rather than replacing human cognition. Based on SCT, AI can be conceptualised as an effective mediational tool that reshapes how learners approach complicated academic tasks. By reducing lower level linguistic concerns, such as grammar and vocabulary, students can concentrate more on cognition-requiring activities, namely idea development and argumentation.

While prior studies have described AI as a form of cognitive support (Kasneci et al., 2023; Dwivedi et al., 2023), the findings of this study suggest a more context-specific role. In EFL settings like in Vietnam, AI appears to be increasingly necessary, as students depend on it for academic lexical resources that are beyond their reach. This finding implies that AI not simply improves learner performance, but also facilitates their participation in academic discourse. Thus, AI functions not merely as a support tool but as a mediational scaffold that expands learners' capacity to participate within their ZPD. Therefore, pedagogically, instead of discouraging AI use entirely, lecturers may adopt structured activities involving brainstorming and outlining with AI assistance into early writing stages, which may help novice writers engage more effectively with academic writing tasks. Especially for first-year students who frequently struggle to generate and organise ideas, AI may play an important role in offering guided pre-writing resources. However, rather than simply copying AI-generated suggestions, students should be required to revise, improve and further develop them with their own thinking, so that they remain actively and critically engaged in the idea generating process.

All in all, AI is not just a powerful tool for improvement, but also a condition that allows learners to engage in academic writing in linguistically constrained contexts.

5.2. The Appropriation of AI Tools: From Dependence to Critical Engagement

The shift from early dependence on AI to more strategic use depicts a dynamic scaffolding process within the ZPD. In fact, students initially heavily depend on AI-generated content due to limited confidence and language ability, which reflects a large amount of external support from this technology. At this stage, AI proves to be a powerful tool that enables students to complete tasks beyond their ability. However, over time, this dependence gradually reduces when students begin to allocate the use of AI for more selective purposes, namely evaluating logical coherence or refining structures. This change indicates a gradual fading of scaffolding and increased self-regulation.

These findings question the concerns that the use of AI may lead to long-term dependency (Kasneci et al., 2023). Instead, these findings can be better understood as a developmental phase, where scaffolding is initially necessary but slowly and progressively fades as students gain competence. This interpretation can also be observed in recent studies about the student engagement with AI in education (Cañete et al., 2026; Park, 2024). This finding suggests that AI-mediated dependence should not be viewed as an inevitable and fixed outcome but instead as a transitional stage within a scaffolding process that supports students developing their self-regulated writing skills.

Based on the findings, the author suggests that instructors should set more focus on teaching their students how to critically integrate AI into their writing skills, rather than simply permitting or restricting AI use. For instance, lecturers can ask students to work in groups to analyse AI-generated texts, identify weaknesses in content or reasoning in the beginning and finally revise the drafts using their own ideas and knowledge. In this way, students may srengthen their analytical skills and promote their active engagement in the writing process, and thereby avoid excessive reliance on AI-generated content.

On the whole, lecurers' directions on appropriate AI usage may help students take advantage of this modern technology to develop their writing skills without becoming over-reliant on it.

5.3. Psychological Support and the EFL Writing Context

In addition to cognitive support, AI also significantly influences students' emotional experiences in writing. Students' reported increased confidence and reduced anxiety indicate that AI serves as an affective scaffolding that supports learners in both task performance and emotional regulation.

When viewed through the lens of SCT, especially in EFL settings, concerns about linguistic accuracy can be interpreted as a breakdown in mediation, where language acting as a mediational tool fails to support cognitive activities. In other words, writing anxiety can limit students' ability to perform effectively within the ZPD, preventing them from engaging with complex ideas. By providing immediate feedback and support, AI reduces emotional barriers, allowing students to actively participate in complex academic writing. As a result, AI may expand the ZPD by lowering both cognitive and emotional difficulties.

While prior studies have emphasised improvements in writing quality by AI (Baidoo-Anu & Ansah, 2023; Delikoura et al., 2025; Kasneci et al., 2023), this study suggests a deeper and more fundamental change in how learners experience writing. Nonetheless, this result also raises another concern: If students' confidence depends largely on AI, they may develop a reliance on external scaffolding both cognitively and affectively.

According to the findings, writing teachers may integrate AI-supported activities into introductory academic writing phases. For example, students can be asked to use AI to generate an outline or revise a draft and then reflect on how they modified the AI-generated suggestions. Indeed, these activities may help reduce writing anxiety and increase confidence while encouraging active engagement in the writing process. However, after the outlining sections, students should be required to independently complete writing tasks using their modified suggestions to prevent over-reliance on AI.

In short, while AI plays a significant role in expanding the ZPD by lowering both cognitive and emotional barriers, it may also cause students to rely on external support for their confidence, rather than on their own competence.

5.4. Authorship, Agency, and Ethical Boundary-Making

Despite frequent use of AI, students continue to position themselves as the main authors. This finding suggests that AI does not reduce their control, but changes how they manage the writing process, as students themselves decide when and how AI is used.

Setting ethical boundaries (e.g. Students reported not using AI in reflective writing tasks) reflects their self-regulation within a mediated context. Students actively decide when scaffolding by AI is necessary and when it should be withdrawn, showing awareness of their ZPD and responsibility for their writing.

At the same time, the combination of frequent AI use and strong claims of authorship implies how students redefine text ownership. Accordingly, authorship is no longer understood as working independently, but as controlling how AI is used and how meaning is made.

This finding is in line with recent studies about concerns regarding academic integrity in AI-assisted writing (Cotton et al., 2023; Chan & Hu, 2023). Furthermore, this result offers a more nuanced interpretation that students actively shape AI's role in their writing outputs, rather than passively relying on it.

The findings highlight the need for clearer guidelines on AI use in academic writing. In paricular, instructors can provide students with concrete examples of acceptable and unacceptable AI practices. For example, students may be allowed to use AI for brainstorming, outlining and language revision, but not for generating ideas or substantial sections of an assignment. This may help students make ethical decisions about AI use, and at the same time ensuring that the ideas expressed in their writing remain original. Not only in class, it is necessary that universities and institutions establish clear policies on the responsible use of AI in academic writing. Such methods should help recognise the ethical boundaries and raise awareness about maintaining their ownership.

As a result, learner agency is not reduced in AI-mediated writing but transformed into the ability to manage AI use and its support within the ZPD.

6. Conclusion

This research offers a deep insight into how English-major students at HUFLIT experience using AI-based tools in academic writing. Based on the qualitative data from 18 participants, the findings show that AI has become an indispensable part in students' academic writing practices in an EFL setting.

First, students consistently perceive AI as a multifaceted writing assistant that facilitates idea generation, organization and language refinement. Second, the research also reveals a transition from early reliance to more strategic and selective use of AI, suggesting that students develop self-regulation within ZPD. Third, besides cognitive support, AI functions as an affective reliever that reduces anxiety and enhances confidence in academic writing activities. Last but not least, students show a strong sense of authorship and actively establish ethical boundaries to maintain academic integrity.

These findings theoretically contribute to SCT by demonstrating that AI-mediated writing is not only about improving performance but also functions as a developmental process shaped by mediation and scaffolding. Practically, the findings suggest that AI should be integrated into EFL writing context as a supportive tool. To put it differently, institutions should guide students to use AI strategically, especially for prompting and critical feedback. Although the findings indicate that AI reduces anxiety and hence improves academic

performance, its use should be gradually withdrawn to promote self-regulation within the ZPD. In addition, universities are advised to establish clear and context-sensitive guidelines to ensure learners' integrity. Overall, the primary challenge is to integrate AI in educational context appropriately to achieve learning benefits and maintain academic integrity.

This study is limited by several issues. First of all, its small sample size (18 participants) and its focus on one single context at HUFLIT may prevent the findings from being transferred to other educational settings. Besides, the aim of qualitative design is to provide in-depth insights into participants' experiences rather than broad generalisation. Next, the cross-sectional design captures students' perceptions at a single point in time; therefore, the findings might not fully reflect how their engagement with AI evolves. Further research should apply longitudinal designs and include a broader range of participants from various institutions in Vietnam to explore how students' pattern of AI use and their perspectives towards this phenomenon develop over time and ensure the richness of findings.

In conclusion, students' engagement with AI in writing practices demonstrates a balance between technological support and individual agency, emphasising the importance of guiding students how to use AI in academic writing responsibly and strategically.

References

Alhusaiyan, E. (2024). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies*. https://doi.org/10.1108/sjls-07-2024-0039

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4337484

Barzanji, A. M., & Alsofyani, A. H. (2026). Exploring Saudi EFL learners' motivation and perceptions of ChatGPT and instructor feedback in paragraph writing. Arab World English Journal, 5-25. https://doi.org/10.24093/awej/ai3.1

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology. https://doi.org/10.1191/1478088706qp063oa

Cañete, P. A. V., Albarillo, R. A., Agsaullo, R. T., Bucayong, J. A., Cadano, M. A., Carajay, M. J. D., Dublin, J. A. M., Exilisi, P. D. A., Estepa, S. J. L., Hilot, L. R. A., Lallen, E. P., Marquez, R. B., Marquez, J. C. S., Rulona, R. M., Sumalig, J. J. A., Tuazon, M. J. A., Pelandas, A. M. O., & Labadan, J. T. (2026). The influence of artificial intelligence usage and learner autonomy on learning engagement among general academic strand students. International Journal of Research and Innovation in Social Science, 10(2), 4282-4294. https://doi.org/10.47772/ijriss.2026.10200313

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1). https://doi.org/10.1186/s41239-023-00411-8

Chanpradit, T. (2025). Generative artificial intelligence in academic writing in higher education: A systematic review. *Edelweiss Applied Science and Technology*, *9*(4), 889–906. https://doi.org/10.55214/25768484.v9i4.6128

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International. https://doi.org/10.1080/14703297.2023.2190148

Delikoura, I., Fung, Y. R., & Hui, P. (2025). From superficial outputs to superficial learning: Risks of large language models in education. arXiv preprint arXiv:2509.21972. https://doi.org/10.48550/arxiv.2509.21972

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., & Carter, L. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of chatgpt: implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460-474. https://doi.org/10.1080/14703297.2023.2195846

Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.58680/ccc198115885

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are enough? an Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Hyland, K. (2019). *Second Language Writing*. Cambridge University Press.

Jacob, S. R., Tate, T., & Warschauer, M. (2024). Emergent AI-assisted discourse: a case study of a second language writer authoring with ChatGPT. *Journal of China Computer-Assisted Language Learning*. https://doi.org/10.1515/jccall-2024-0011

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Stadler, M. (2023). Chatgpt for good? on opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Kvale, S., & Brinkmann, S. (2015). *Interviews*. Sage.

Park, S. (2024). Transition in the concept of scaffolding. In Massive Open Online Courses - Learning Frontiers and Novel Innovations. IntechOpen. https://doi.org/10.5772/intechopen.1006202

Patton, M. Q. (2014). Qualitative Research & Evaluation Methods: Integrating Theory and Practice. In *Google Books*. SAGE Publications.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Sanz-Tejeda, A., Domínguez-Oller, J. C., Baldaquí-Escandell, J. M., Gómez-Díaz, R., & García-Rodríguez, A. (2026). The impact of generative AI on academic reading and writing: a synthesis of recent evidence (2023–2025). *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1711718

Thorne, S. L., & Lantolf, J. P. (2006). A linguistics of communicative activity. In Disinventing and Reconstituting Languages (pp. 170-195). Multilingual Matters. https://doi.org/10.2307/jj.27939678.11

Vygotsky, L. S. (1980). Mind in society. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Warschauer, M. (2010). New tools for teaching writing. *Language Learning & Technology*, *14*(1), 3–8. https://doi.org/10.64152/10125/44196

Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*, 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zhai, X. (2023). ChatGPT for Next Generation Science Learning. *XRDS: Crossroads, the ACM Magazine for Students*, *29*(3), 42–46. https://doi.org/10.1145/3589649

Zhao, X., Cox, A., & Cai, L. (2024). ChatGPT and the digitisation of writing. *Humanities & Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-02904-x

Bionote

Phan Ngoc Phuong Vy received her Bachelor's degree in English Language Teacher Education from Ho Chi Minh City University of Education and her Master's degree in Applied Linguistics from Curtin University. She is currently a lecturer at the Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT). Her academic interests include English language pedagogy, language teaching methodology, academic writing, and technology-enhanced language learning. Her research focuses on teaching and learning practices in English language education, particularly the integration of digital tools and emerging technologies to support student learning.

Appendices

Appendix A: Semi-structured Interview Questions

1. Can you tell me about your experience with academic writing at university?
   - What kinds of writing tasks do you usually do?
2. Can you describe how you first started using AI tools for academic writing?
   - What motivated you to try them?
   - Who or what influenced your decision?
3. Can you walk me through how you typically use AI in a writing assignment?
   - At what stage do you usually use it? (brainstorming, drafting, revising, proofreading)
4. Can you describe a recent writing task where you used AI?
   - What did you ask the AI to do?
   - What did you do with the AI-generated content?
5. How do you usually interact with AI tools when writing?
   - Do you modify your prompts?
   - Do you critically evaluate or edit the AI's responses?
6. Has your way of using AI changed over time? Why?
- What has changed?
7. Are there situations where you choose not to use AI? Why?
8. How do you see the role of AI in your academic writing?
   - Do you see it as an assistant, collaborator, shortcut, or something else?
9. In what ways do you think AI shapes or transforms your approach to academic writing?
   - Does it influence your ideas, structure, language, or confidence?
10. How do you feel about relying on AI for writing tasks?
    - Do you ever feel dependent on it?
11. What concerns, if any, do you have about using AI in academic writing?
    - Academic integrity?
    - Skill development?
    - Fear of detection?
12. When using AI, who do you feel is the main author of the text? Why?
13. How do you think using AI affects your development as a writer?
    - Do you feel it helps you learn, or does it reduce your learning?
14. Is there anything else about your experience using AI for academic writing that you would like to share?

#21 EFL-TO-ESL REPOSITIONING THROUGH TEACHER DEVELOPMENT: IMPLICATIONS FOR INTERNATIONALIZED EDUCATION AND WORKPLACES

Nguyễn Hưng Bình[1]

Abstract

Amid accelerating globalization, English is increasingly repositioned from a foreign language to a functional second language, yet the effectiveness of teacher professional development in enabling this shift remains uneven. This study investigates how professional development influences EFL-to-ESL repositioning and its alignment with internationalized education and workplace demands across four high schools in Vietnam and the Philippines. Adopting a mixed-methods, multi-site case study design, the research integrates survey data, interviews, and classroom observations. Findings reveal that sustained, context-sensitive professional development significantly enhances communicative pedagogy, fostering stronger alignment with real-world language use. Institutional disparities highlight the role of resources and professional learning cultures in shaping outcomes. The study contributes empirically grounded insights into scalable teacher development models, informing policy and practice in internationalizing education systems.

Keywords: alignment, internationalization, professional development, repositioning

TÁI ĐỊNH VỊ VAI TRÒ TỪ EFL ĐẾN ESL THÔNG QUA PHÁT TRIỂN ĐỘI NGŨ GIÁO VIÊN: HÀM Ý QUỐC TẾ HOÁ GIÁO DỤC VÀ NHIỆM SỞ

Tóm tắt

Trong bối cảnh toàn cầu hóa đang gia tăng nhanh chóng, tiếng Anh ngày càng được định vị lại từ một ngoại ngữ thành một ngôn ngữ thứ hai mang tính chức năng, tuy nhiên hiệu quả của hoạt động phát triển chuyên môn giáo viên trong việc thúc đẩy sự chuyển đổi này vẫn chưa đồng đều. Nghiên cứu này khảo sát cách thức phát triển chuyên môn ảnh hưởng đến quá trình định vị lại từ EFL sang ESL và sự phù hợp của quá trình này với các yêu cầu quốc tế hoá giáo dục và nhiệm sở tại bốn trường trung học ở Việt Nam và Philippines. Áp dụng thiết kế nghiên cứu tình huống đa địa điểm theo phương pháp hỗn hợp, nghiên cứu kết hợp dữ liệu khảo sát, phỏng vấn và quan sát lớp học. Kết quả cho thấy hoạt động phát triển chuyên môn được duy trì và phù hợp với bối cảnh đã nâng cao đáng kể phương pháp giảng dạy giao tiếp, thúc đẩy sự gắn kết chặt chẽ hơn với việc sử dụng ngôn ngữ trong thực tế. Những khác biệt giữa các cơ sở giáo dục làm nổi bật vai trò của nguồn lực và văn hóa học tập chuyên môn trong việc định hình kết quả. Nghiên cứu đóng góp những hiểu biết có cơ sở thực nghiệm về các mô hình phát triển giáo viên có khả năng mở rộng, qua đó cung cấp cơ sở cho chính sách và thực tiễn trong quá trình quốc tế hóa các hệ thống giáo dục.

Từ khóa**:** sự phù hợp, quốc tế hóa, phát triển chuyên môn, định vị lại

[1] *MA student, Phenika University, Hanoi, Vietnam; Email: thtam@apd.edu.vn*

1. Introduction

1.1. Background:

Across rapidly globalizing economies, English is increasingly repositioned from a foreign language to a functional second language embedded in education and work. In Southeast Asia, this shift is particularly visible as schools respond to internationalization agendas and evolving workplace language demands that prioritize communicative competence, intercultural awareness, and digital fluency. In Vietnam and the Philippines, English occupies distinct yet converging roles within national curricula, creating a fertile ground to examine how pedagogical transformation unfolds in practice. Teacher professional development has emerged as a central lever in this transition, equipping educators with the capacity to integrate learner-centered approaches, technology-enhanced instruction, and contextually relevant language use. However, the extent to which such development initiatives effectively support the EFL-to-ESL repositioning remains uneven and underexplored. The selection of Phenikaa High School (PHS) in Hanoi and Tam Duong High School (TDHS) in Phu Tho Province reflects contrasting urban–rural contexts within Vietnam, while Far Eastern High School (FEHS) and Philippine Normal University High School (PNHS) represent established English-medium traditions in the Philippines. Together, these institutions provide a comparative lens to investigate how differing policy environments, resource availability, and professional learning cultures shape the enactment of English language repositioning in secondary education.

1.2. Aims:

Despite growing policy attention, a critical gap persists in understanding how teacher professional development concretely influences pedagogical shifts toward ESL-oriented practices and how these shifts correspond with internationalized education and workplace expectations. This study addresses this gap by examining two core questions: how teacher professional development influences EFL-to-ESL repositioning across selected high schools, and to what extent such repositioning aligns with broader global and occupational language demands. The research aims to identify key dimensions of effective professional development, analyze changes in teachers' instructional practices and beliefs, and evaluate the alignment between classroom language use and real-world communicative requirements. It is hypothesized that sustained, context-sensitive professional development significantly enhances ESL-oriented pedagogy and fosters stronger alignment with international and workplace standards. The study holds significance for policymakers, educators, and curriculum designers by offering empirically grounded insights into scalable models of teacher learning and language policy implementation. Ultimately, by bridging institutional practices with global demands, this research seeks to inform more coherent strategies for English language education reform. These considerations collectively justify the necessity of conducting this multi-site case study in both Vietnam and the Philippines.

1.3. Method

This study adopts a mixed-methods, multi-site case study design to examine how teacher professional development shapes EFL-to-ESL repositioning and its alignment with internationalized education and workplace language demands. A convergent research framework integrates quantitative survey data with qualitative interview and classroom observation evidence, enabling triangulation across four institutions: PHS and TDHS in Vietnam, and FEHS and PNHS in the Philippines. Data were collected over twelve weeks through online platforms such as Google Forms, Zoom, and Microsoft Teams. The sample

comprises approximately 80 English teachers and 200 senior high school students, selected through purposive and stratified sampling to ensure contextual diversity. Instruments include structured questionnaires measuring pedagogical beliefs and practices, semi-structured interviews exploring professional development experiences, and observation checklists assessing ESL-oriented instructional behaviors. Quantitative data are analyzed using descriptive and inferential statistics, including regression analysis to determine the impact of professional development on pedagogical shifts. Qualitative data are coded thematically to capture patterns in teacher cognition and classroom enactment. Instrument reliability is established through pilot testing and Cronbach's alpha, while validity is ensured via expert review and data triangulation. All participants provided informed consent, and participation was entirely voluntary. To safeguard confidentiality, all data were anonymized and encrypted, reinforcing ethical compliance and enhancing the credibility and transparency of the findings.

2. Literature review

Grounded in sociocultural and professional identity frameworks, this review synthesizes prior scholarship to contextualize EFL-to-ESL repositioning through teacher development. Studies were purposively selected based on relevance to teacher identity, professional learning, and alignment with global language demands across Asian contexts. Both empirical and theoretical works employing qualitative, quantitative, and mixed-methods designs are examined. Analytical focus is placed on conceptual definitions, methodological rigor, and reported impacts on pedagogy, thereby establishing a coherent foundation for comparative, multi-site investigation in this study.

Findings from Aultan, A. et al. (2020), Malaysia-based, investigated how sociocultural environments shape EFL-to-ESL transitions and teachers' professional identity. Using qualitative interviews and narrative inquiry, the study conceptualizes identity as socially constructed and context-dependent. Findings reveal that exposure to ESL environments enhances teachers' adaptability and communicative orientation. However, limited institutional support constrains identity transformation. While offering valuable sociocultural insights, the study lacks cross-national comparison and quantitative validation. This gap justifies examining how structured professional development systematically supports repositioning across varied institutional contexts in the present research. Besides, Ling, C. (2021), conducted in China, explores emotional dimensions in EFL/ESL teachers' identity development through a psychological lens. Employing survey and interview methods, the study defines emotions as mediators of professional growth. Results indicate that positive emotional engagement strengthens pedagogical commitment and identity reconstruction. Nonetheless, the research focuses primarily on internal factors, overlooking structural influences such as training systems. Its limitation lies in the absence of linkage between emotions and instructional outcomes, thereby necessitating further investigation into how professional development integrates affective and pedagogical dimensions. Echoing earlier studies, Xiaodong, L. (2022) provides a comprehensive theoretical review of EFL/ESL teacher professional development, synthesizing diverse approaches including reflective practice, mentoring, and technology integration. The study conceptualizes professional development as a continuous, multidimensional process influencing teaching quality. While it outlines significant impacts on pedagogy, it remains largely conceptual without empirical validation. Additionally, the review does not explicitly address alignment with workplace demands. This limitation underscores the need for data-driven studies examining how professional development translates into measurable ESL-oriented practices in real educational settings.

In their analysis, Sayit, A. et al. (2025), based in Indonesia, employed a quantitative design to assess the impact of teacher professional development on instructional practices and student achievement. Using statistical analysis, the study demonstrates a strong positive relationship between training and improved teaching outcomes. However, its focus remains within EFL contexts, without addressing ESL repositioning or global alignment. Moreover, contextual variables are minimally explored. These constraints highlight the necessity for broader comparative research linking professional development to both pedagogical transformation and internationalized language demands. Similarly, Xiaolin, L. (2023) theorized the interplay among professional identity, agency, and positioning within EFL contexts. Drawing on discourse analysis and conceptual synthesis, the study emphasizes teachers' active role in negotiating pedagogical change. While it advances theoretical understanding, empirical application is limited. The absence of concrete classroom evidence and cross-contextual comparison restricts its practical implications. Consequently, further research is required to operationalize these constructs within real-world ESL transitions, particularly through structured professional development initiatives examined in this study. Moreover, Yusnamariah, M. et al. (2025), Malaysia, analyze current trends in English teacher professional development through literature review and survey data. The study identifies shifts toward digital integration, collaborative learning, and lifelong professional growth. Despite its breadth, it lacks focus on EFL-to-ESL repositioning and does not examine workplace alignment. Furthermore, contextual disparities across institutions are underexplored. These gaps reinforce the need for targeted empirical studies investigating how contemporary professional development models influence ESL-oriented pedagogy and global communicative competence.

In a study conducted by Nguyen Van Thong et al. (2023) in Vietnam, they examined lesson study as a model for teacher development using qualitative case study methods. Findings highlight reflective practice and peer collaboration as catalysts for pedagogical improvement. However, the study is limited to a single national context and focuses primarily on instructional reflection rather than ESL repositioning. Its narrow scope suggests the necessity for comparative, multi-site research linking professional development approaches to broader internationalization and workplace language outcomes. Furthermore, Cong Khanh Ly (2022), Vietnam, investigated learners' beliefs about English as a global language using survey methodology. The study defines English as a tool for global communication and mobility. Findings reveal increasing recognition of its practical value among students. Nevertheless, the research does not address teacher development or pedagogical transformation. This limitation indicates a gap in connecting learner perceptions with instructional practices, supporting the rationale for examining teacher-driven ESL repositioning in the present study. Last but not least, Emma, D. (2021) explored English-Medium Instruction and teacher development within internationalization agendas through survey-based research in higher education. The study highlights the importance of training in enhancing instructional effectiveness and global alignment. However, it focuses on tertiary contexts and does not address secondary education or EFL-to-ESL transitions. Additionally, workplace alignment is implied rather than empirically measured. These gaps justify extending the investigation to high school settings and explicitly examining alignment with real-world language demands.

Collectively, prior studies emphasize teacher identity, professional development, and global English usage, yet remain fragmented across theoretical, emotional, and pedagogical dimensions. Few integrate these aspects within a coherent framework linking professional development to ESL repositioning and workplace alignment. Empirical evidence is often context-specific, lacking cross-national comparison and multi-site validation. Moreover,

alignment with real-world communicative demands is insufficiently operationalized. Addressing these gaps, the present study investigates how structured professional development drives EFL-to-ESL repositioning and its alignment with internationalized education and workplace expectations across diverse institutional contexts, offering novel comparative and practice-oriented contributions.

3. Results and findings

3.1. Quantitative data analysis

The results section presents quantitative and qualitative evidence examining the influence of teacher professional development on EFL-to-ESL repositioning and its alignment with internationalized education and workplace language demands across four high schools in Vietnam and the Philippines.

Table 1. Professional development impact on ESL repositioning

School	Mean score (%)
PHS	78
TDHS	65
FEHS	85
PNHS	88

(Source: Author, 2026)

Table 1 indicates clear cross-site variation in the impact of teacher professional development on EFL-to-ESL repositioning. PNHS records the highest mean (88%), followed by FEHS (85%), demonstrating strong ESL-oriented pedagogical transformation aligned with communicative and workplace demands. In contrast, PHS shows moderate impact (78%), while TDHS reports the lowest (65%), suggesting limited repositioning capacity in less-resourced contexts. These findings confirm that sustained, context-sensitive professional development significantly enhances ESL practices and better aligns classroom instruction with internationalized education and real-world language use expectations. Descriptive statistics reveal noticeable variation in the perceived impact of teacher professional development across institutions. As illustrated, mean scores for professional development impact on ESL-oriented pedagogy range from 65 at TDHS to 88 at PNHS, indicating stronger pedagogical transformation in Philippine contexts.

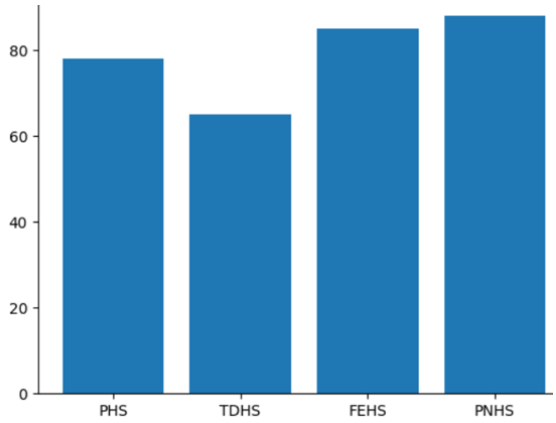


Figure 1: Personal development impact on ESL repositioning (Source: Author, 2026)

Figure 1 illustrates notable disparities in the impact of professional development (PD) on ESL repositioning across the four schools. PNHS achieves the highest score (88%), closely followed by FEHS (85%), indicating the strong effectiveness of PD in fostering communicative, ESL-oriented practices. PHS demonstrates a moderate level (78%), suggesting partial alignment with internationalized education demands. In contrast, TDHS records the lowest score (65%), reflecting constraints in translating PD into practice. Overall, the pattern confirms that sustained and context-responsive PD significantly enhances pedagogical shifts toward ESL and workplace-relevant language use. The data indicate that schools with sustained and structured professional development programs demonstrate higher levels of ESL-oriented instructional practices. Classroom observations confirm that teachers in FEHS and PNHS more frequently employed communicative tasks, authentic materials, and interaction-based techniques, whereas TDHS exhibited more traditional, grammar-focused instruction.

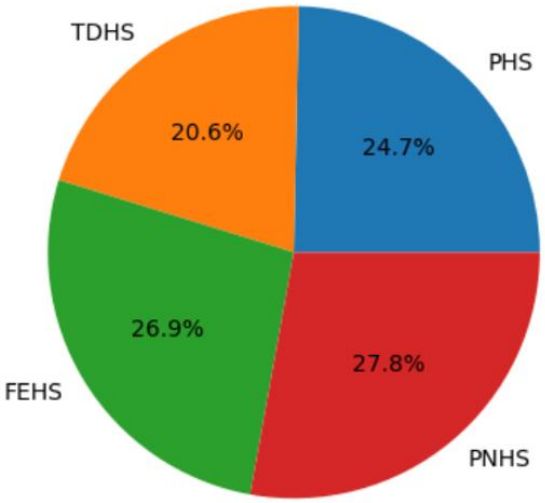


Figure 2: Distribution of PD impact (Source: Author, 2026)

Figure 2 presents the proportional distribution of professional development (PD) impact on ESL repositioning across the four schools. PNHS accounts for the largest share (27.8%), followed by FEHS (26.9%), highlighting the Philippines' stronger alignment with ESL-oriented practices. PHS contributes 24.7%, indicating moderate influence, while TDHS represents the smallest proportion (20.6%), reflecting comparatively limited PD effectiveness. This distribution reinforces that more robust and sustained PD frameworks are associated with greater pedagogical transformation and closer alignment with internationalized education and workplace communication demands. The chart demonstrates that PNHS and FEHS together account for over half of the total impact, reflecting their institutional emphasis on continuous teacher training and internationalized curricula. In contrast, TDHS contributes the smallest proportion, highlighting disparities in access to professional learning opportunities.

Table 2. Alignment with internationalized education and workplace demands

School	Alignment Score (%)
PHS	70
TDHS	60
FEHS	82
PNHS	86

(Source: Author, 2026)

Table 2 reveals clear differences in the degree of alignment between ESL repositioning and internationalized education and workplace demands across the four schools. PNHS records the highest alignment score (86%), followed by FEHS (82%), indicating strong responsiveness to global communication standards. PHS shows a moderate level (70%), suggesting partial integration of workplace-oriented language practices. In contrast, TDHS reports the lowest score (60%), reflecting limited alignment. These findings suggest that stronger professional development contributes significantly to enhancing alignment with international and occupational language requirements. Further analysis examines alignment with internationalized education and workplace language demands. As presented, alignment scores follow a similar pattern, with PNHS (86) and FEHS (82) outperforming their Vietnamese counterparts.

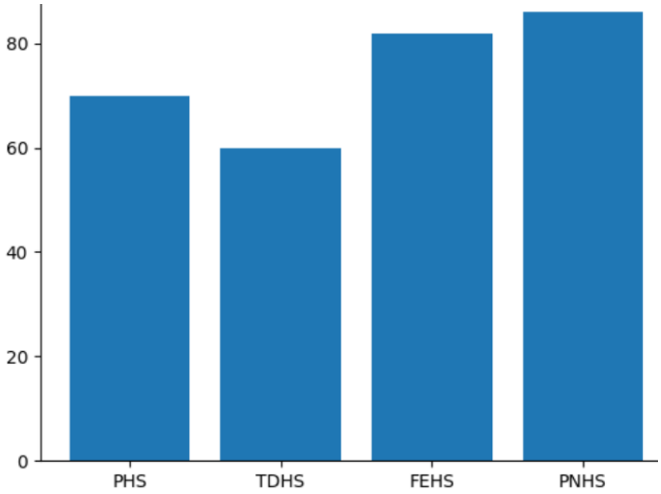


Figure 3: Alignment with workplace demands (Source: Author, 2026)

Figure 3 demonstrates varying levels of alignment between ESL repositioning and workplace language demands across the four schools. PNHS achieves the highest score (87%), followed by FEHS (82%), indicating strong integration of communicative and job-relevant language skills. PHS records a moderate level (70%), suggesting partial alignment with international workplace expectations. In contrast, TDHS shows the lowest score (60%), reflecting limited capacity to connect classroom practices with real-world language use. Overall, the results highlight that stronger professional development correlates with better workplace-oriented ESL alignment.

Inferential statistical analysis further supports these observations. Regression results indicate a strong positive relationship between teacher professional development and ESL repositioning ($\beta = 0.72$, $p < 0.01$), suggesting that professional learning significantly predicts pedagogical transformation. Similarly, professional development is moderately correlated with alignment to workplace language demands ($r = 0.65$), indicating that improved teaching practices contribute to more relevant language use.

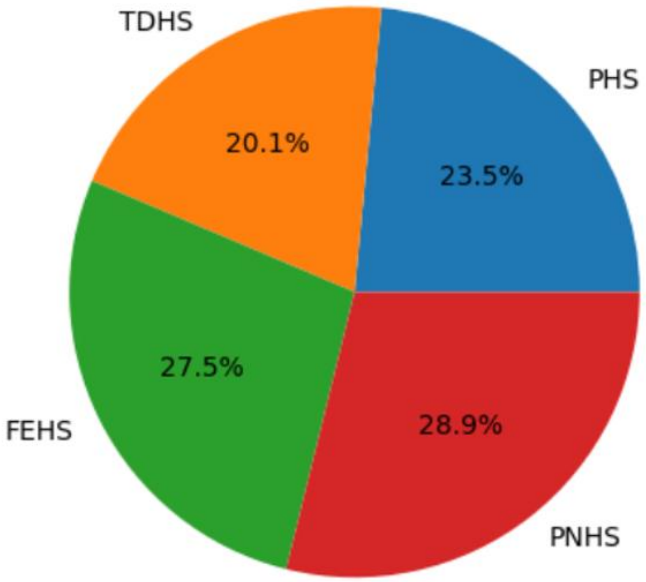


Figure 4: Alignment distribution (Author, 2026)

Figure 4 shows the distribution of alignment between ESL repositioning and workplace language demands across the four schools. PNHS accounts for the largest proportion (28.9%), followed by FEHS (27.5%), indicating stronger responsiveness to global workplace communication needs. PHS represents 23.5%, reflecting moderate alignment, while TDHS contributes the smallest share (20.1%), suggesting limited integration of practical language use. Overall, the distribution confirms that higher levels of professional development are associated with more effective alignment with internationalized education and workplace expectations. This visualization confirms that Philippine institutions demonstrate stronger alignment with global communication requirements, likely due to their longer-standing integration of English as a medium of instruction.

3.2. Qualitative data analysis

Qualitative findings from interviews provide further insight into these patterns. Teachers in high-performing schools reported increased confidence in facilitating communicative learning environments and integrating technology into instruction. They emphasized the role of ongoing mentoring, peer collaboration, and reflective practice in sustaining pedagogical change. Conversely, teachers in lower-performing contexts cited limited training duration, lack of institutional support, and curriculum constraints as barriers to effective ESL repositioning. Student responses also reflect these differences. Learners in FEHS and PNHS reported higher engagement in speaking and collaborative tasks, as well as greater exposure to real-world language scenarios. In contrast, students in TDHS indicated a stronger focus on exam preparation and limited opportunities for authentic communication.

Overall, the findings demonstrate that teacher professional development plays a significant role in facilitating the transition from EFL to ESL across diverse educational contexts. The data consistently show that schools with more comprehensive and sustained professional development initiatives achieve higher levels of pedagogical transformation and stronger alignment with international and workplace language demands. From a Vietnam-focused perspective, the results highlight the need to strengthen teacher development systems, particularly in rural or resource-constrained settings. While PHS demonstrates promising progress, the gap between Vietnamese and Philippine institutions suggests that further

investment in training quality, duration, and contextual relevance is necessary. Importantly, the study underscores that professional development must move beyond theoretical instruction toward practical, classroom-based application to achieve meaningful ESL repositioning.

In summary, the findings directly address the research questions by demonstrating that teacher professional development significantly influences EFL-to-ESL repositioning and that such repositioning is positively associated with alignment to internationalized education and workplace communication needs. These results provide robust empirical support for the study's hypothesis and offer actionable insights for educational policy and practice.

4. Conclusion

Findings from this multi-site investigation substantively clarify how teacher professional development shapes the repositioning of English from an EFL orientation toward an ESL paradigm across contrasting educational contexts. The evidence indicates that sustained, context-responsive professional learning fosters a decisive shift from form-focused instruction to communicative, interaction-driven pedagogy. Teachers who engaged in iterative training, mentoring, and reflective practice demonstrated greater capacity to integrate authentic materials, digital tools, and task-based activities, thereby aligning classroom discourse with real-world language use. Conversely, limited or fragmented development constrained pedagogical transformation, maintaining exam-oriented practices and restricting learners' communicative exposure. These patterns directly answer the first research question by confirming that the depth, continuity, and contextualization of professional development determine the extent of EFL-to-ESL repositioning. Furthermore, the convergence between greater pedagogical change and higher alignment with internationalized education and workplace demands addresses the second question, suggesting that ESL-oriented instruction operates as a mediating mechanism linking teacher learning to employability-relevant competencies. Theoretically, these findings extend models of language policy enactment by foregrounding teacher agency as a critical driver of systemic change. Practically, they underscore the necessity of long-term, practice-embedded development frameworks capable of translating policy aspirations into classroom realities. Collectively, the study contributes a nuanced comparative perspective, highlighting how institutional conditions and professional cultures shape the trajectory of English language transformation in internationalizing education systems.

Limitations: Despite its contributions, the study is not without limitations. The sample size, while sufficient for exploratory analysis, remains relatively modest and may limit generalizability beyond the selected institutions. Contextual differences between urban and rural settings, as well as national policy environments, may also introduce variability not fully captured in the design. Additionally, reliance on self-reported data alongside observations may involve response bias. The cross-sectional timeframe further restricts the ability to assess the long-term sustainability of pedagogical change.

Recommendations: Future research should adopt longitudinal designs to examine the durability of EFL-to-ESL repositioning over time and across broader institutional contexts. Expanding the sample to include additional regions and school types would enhance comparative validity. Practically, education systems should prioritize sustained, context-sensitive professional development models that integrate mentoring, peer collaboration, and classroom-based application. Policymakers are encouraged to align teacher training with evolving workplace language demands, ensuring that instructional practices remain relevant, adaptive, and responsive to global communicative expectations.

Reference

Aultan, A., et al. (2020). From EFL to ESL context: The impact of sociocultural environment on teachers' professional identity development. *3L: The Southeast Asian Journal of English Language Studies, 26*(4), 74–86. https://doi.org/10.17576/3L-2020-2604-06

Cong Khanh Ly. (2022). English as a global language: An exploration of EFL learners' beliefs in Vietnam. *International Journal of TESOL & Education, 3*(1), 19–33. https://doi.org/10.54855/ijte.23312

Emma, D. (2021). Repositioning English-medium instruction in a broader international agenda: Insights from a survey on teacher professional development. *Alicante Journal of English Studies, 34*, 15–38. https://doi.org/10.14198/raei.2021.34.08

Ling, C. (2021). The implications of EFL/ESL teachers' emotions in their professional identity development. *Frontiers in Psychology, 12*, Article 755592. https://doi.org/10.3389/fpsyg.2021.755592

Nguyen Van Thong, et al. (2023). Using lesson study for teacher development: A case study of Vietnamese EFL teachers' reflections. *CTU Journal of Innovation and Sustainable Development, 15*(3), 72–80. https://doi.org/10.22144/ctujoisd.2023.052

Sayit, A., et al. (2025). The impact of teacher professional development on EFL teachers' instructional practices and students' academic achievement. *Ethical Lingua: Journal of Language Teaching and Literature, 12*(2), 668–683. https://doi.org/10.30605/25409190.877

Xiaodong, L. (2022). A theoretical review on EFL/ESL teachers' professional development: Approaches, applications, and impacts. *Frontiers in Psychology, 13*, Article 912365. https://doi.org/10.3389/fpsyg.2022.912365

Xiaolin, L. (2023). A theoretical review on the interplay among EFL teachers' professional identity, agency, and positioning. *Heliyon, 9*(4), e15510. https://doi.org/10.1016/j.heliyon.2023.e15510

Yusnamariah, M., et al. (2025). Current trends in the English language teacher professional development. *International Journal of Language Education and Applied Linguistics, 15*(1). https://doi.org/10.15282/ijleal.v15i1.11198

#23 Semiotics and Translanguaging Approaches to Inclusion and Critical Thinking in Generative AI Adoption

Fati Osuman[1]

Abstract

Generative Artificial Intelligence (Gen AI) in second language (L2) education offers new affordances for multilingual interaction, multimodal expression, and individualized feedback. However, concerns around academic integrity, erosion of critical thinking, equity, and the reproduction of cultural and linguistic bias remain underexplored. Existing technology-integration frameworks have been criticized for prioritizing tools over learner agency (Blundell et al., 2022), while recent pedagogical approaches incorporating translanguaging have not fully addressed assessment, reflection, or multimodal task design in AI-supported learning (Donley, 2024). Responding to these gaps, this conceptual paper proposes an inclusive framework for Gen AI–mediated task design grounded in translanguaging pedagogy and social semiotic theory. Drawing on current literature, the paper examines how Gen AI tasks designed through this lens can support three interrelated dimensions of learning: (1) critical thinking; (2) translanguaging and multimodality; and (3) equity and representation. The article concludes with implications and future directions for designing AI-supported language tasks that foster critical AI literacy, inclusion, and meaningful multilingual participation.

Keywords: generative AI, translanguaging, social semiotics, critical thinking, inclusive L2 pedagogy.

1. Introduction

Generative Artificial Intelligence (Gen AI), particularly tools like ChatGPT that generate new content through large-scale language modelling, has gained considerable attention in L2 for its potential to improve teaching and learning efficiency (Khonke & Zou, 2025). Gen AI enables language instructors to rapidly customise learning tasks, scaffold instruction, create multimodal materials, provide instant individualised feedback, and facilitate multilingual interaction in the classroom (Evmenova et al., 2024; Tang & Su, 2024). These affordances are particularly relevant in increasingly multilingual and international educational contexts where learners engage with diverse languages, identities, and modes of communication. While these opportunities open up new possibilities for language learning and feedback, concerns persist around academic integrity, data privacy, and the erosion of teacher autonomy (Crompton et al., 2024). Furthermore, overreliance on Gen AI risks undermining critical thinking development and exacerbating inequities for historically marginalised learners (Evmenova et al., 2024; Ahmed, 2025). These challenges imply that effective AI integration requires more than technological adoption. It thus requires pedagogical approaches that promote learner agency, critical engagement, and inclusion.

In the post-pandemic boom of Gen AI, frameworks like Substitution Augmentation, Modification Redefinition (SAMR) (Puentedura, 2013) guided technology integration but were criticized for prioritizing tool use classification over learner experience and pedagogical context (Blundell et al., 2022), leading scholars such as Donley (2024) to explore options such as translanguaging that centre inclusion and criticality. Donley (2024) used a teacher's qualitative data to show how translanguaging pedagogy guided Gen AI use in language

[1] *University of Ottawa, Canada, Email: fatymaosuman@gmail.com*

classrooms, focusing on three dimensions: affordances, production, and co-labour to promote inclusion and cultural competence. However, the study did not address assessment, reflections and multimodal task design, which are essential to counter overreliance and erosion of critical thinking. Consequently, there remains a need for conceptual frameworks that connect AI use with critical thinking, multilingual meaning-making, and equitable participation.

In response to the gaps identified, this article proposes a conceptual framework that applies translanguaging pedagogy and social semiotic theory to guide inclusive Gen AI-supported language learning. The framework positions learners as active meaning-makers who analyze, critique, and redesign AI-generated content while drawing on their full linguistic, cultural, and semiotic repertoires.

2. Conceptual Foundations

2.1. Translanguaging Pedagogy

It is a pedagogical practice that draws on multilingual learners' full linguistic and cultural resources to make sense of their world and communicate effectively (García, 2014). It fosters culturally responsive learning, reduces learner anxiety, and increases comprehension and participation (Cioè-Peña, 2022). Beyond language switching, translanguaging includes multimodal practices such as reading, writing, and dialogue for meaning-making, emotional expression and identity development (Galante, 2020). García and Kleifgen (2019) propose five dimensions for implementing translanguaging pedagogy in instruction: affordances, production, co-labor, assessment, and reflection. In terms of affordances and production, Gen AI tools can be used to generate multimodal input and support content creation that reflects students' cultural and linguistic identities. Co-labor can be fostered through collaborative, AI-supported learning that validates diverse knowledge systems. In assessment, educators adopt multimodal evaluation procedures that include translanguaging in formative and summative assessments. Reflection, on the other hand, involve examining cross-linguistic comparisons, language usage, and the sociocultural aspects of language to support deeper learning. These dimensions provide useful foundation for designing AI-supported language tasks.

2.2. Social *Semiotic Theory and Multimodality*

Semiotics is the study of signs and how meaning is created through them (Danesi, 2000). Social semiotics, introduced by Gunther Kress, shifts the focus from static sign systems to meaning-making as a dynamic, socially situated practice (Kress, 2010). Central to this approach is multimodality, which sees learning as designing meaning through selecting and reshaping semiotic resources such as images, language, sound, diagrams, and gestures (Adami et al., 2022). These resources carry culturally and socially embedded meanings, offering multiple entry points that make learning more accessible and inclusive. Wanselin et al. (2022) and Cheung et al. (2024) outline three key meaning-making layers relevant to Gen AI adoption in L2 tasks: presentational meaning (ideas, events, and participants), orientational meaning (attitudes, emotions, and perspectives), and organisational meaning (how elements connect to create coherence). Translanguaging and social semiotics combined position learners as analysts (critiquing Gen AI outputs for accuracy, representation, and bias), designers (revising, remixing or creating content), and advocates (producing work that reflects diverse identities). This approach fosters critical AI literacy while maintaining learner agency. When combined with translanguaging pedagogy, it offers a framework for critically evaluating and redesigning AI-generated content while supporting multilingual and multimodal expression.

3. Proposed Framework: Translanguaging and Social Semiotics for Inclusive AI Pedagogy

The proposed Gen AI adoption framework combines translanguaging pedagogy and social semiotic theory to guide inclusive and critical uses of Gen AI in L2 education. It treats AI outputs as semiotic resources that can be examined, critiqued, and transformed through multilingual and multimodal meaning-making. Three core aspects of learning underpin the framework: critical thinking, translanguaging and multimodality, and equity and representation.

3.1. Critical Thinking

A key challenge surrounding Gen AI in education is the potential for learner overreliance on AI-generated responses and limited autonomy. When students accept AI outputs uncritically, they miss opportunities for independent reasoning, evaluation, and knowledge construction. The framework addresses this challenge by positioning learners as critics of AI-generated content rather than passive recipients.

Drawing on the assessment and reflection dimensions of translanguaging pedagogy, learners are encouraged to examine how meaning is constructed in AI-generated texts, images, and other multimodal outputs. Social semiotics provides an analytical lens through which students can evaluate how information is presented (presentational layer), whose perspectives are represented (orientational layer), and how ideas are organized (organizational layer). For example, students might compare AI-generated responses across languages, identify inconsistencies or omissions, and revise outputs to better reflect particular cultural or contextual perspectives. Through these activities, learners develop critical AI literacy while engaging in higher-order processes such as analysis, evaluation, and redesign.

3.2. Translanguaging and Multimodal Meaning-Making

The framework also supports learner agency through multilingual and multimodal meaning-making. Translanguaging pedagogy encourages learners to mobilize their full linguistic and cultural repertoires when constructing knowledge, while social semiotics recognizes that meaning is communicated through multiple modes, including language, images, audio, and visual design.

Within AI-supported tasks, learners assume the role of designers who actively transform and remix content rather than simply consuming it. For example, teachers can use Gen AI to create scaffolded examples that students may use as a starting point for creating multilingual narratives, presentations, podcasts, or digital stories that reflect their experiences and identities. In doing so, they move across languages and modes to communicate meaning in ways that are personally relevant and culturally situated. Such activities extend language learning beyond traditional text-based tasks and create opportunities for creativity, collaboration, and multimodal literacy development.

3.3. Equity and Representation

A third dimension of the framework addresses equity and representation. As Gen AI systems are trained on large datasets that may privilege dominant linguistic and cultural perspectives, AI-generated outputs can reproduce existing biases and marginalize underrepresented groups. To address this, task design should centre cultural representation, inclusivity and accessibility. Through a translanguaging lens, tasks should encouraged

students to draw on diverse linguistic and cultural resources to challenge dominant narratives and contribute alternative viewpoints. Learners may analyze AI-generated texts for cultural assumptions, identify missing perspectives, and redesign content to better reflect their communities and lived experiences. In this way, students become advocates who use AI not only to access information but also to question, reshape, and expand it. Equity also involves ensuring meaningful participation in AI-supported learning. Educators must consider differences in digital access, technological familiarity, and language backgrounds when designing tasks. Providing multiple pathways for engagement and recognizing diverse forms of expression can help ensure that AI-enhanced learning environments remain accessible and inclusive for all learners.

Table 1

Summary of how translanguaging pedagogy and social semiotics can inform AI-supported task design in L2 education.

Translanguaging Dimension	Semiotic Focus	AI-Supported Strategy	Illustrative Task
Affordances	Multimodal input	Generate multilingual and multimodal content	Analyze AI-generated bilingual dialogue for tone and cultural meaning
Production	Presentational meaning	Co-create and adapt content using multiple languages and modes	Develop a multilingual travel blog incorporating visuals and local expressions
Co-labor	Organizational meaning	Support collaborative writing and revision	Co-author a digital story using AI-generated prompts
Assessment	Orientational meaning	Provide formative feedback on multimodal work	Revise an audio-visual reflection using AI feedback and self/peer-assessments
Reflection	Critical meaning-making	Critique and redesign AI-generated content	Rewrite an AI-generated text to address bias or missing perspectives

4. Implications and Conclusion

As Gen AI becomes increasingly adopted in language education, there is a growing call for pedagogical approaches that move beyond technological integration toward critical and inclusive learning. The framework proposed in this article responds to this need by integrating translanguaging pedagogy and social semiotic theory to support critical engagement, multilingual meaning-making, and equitable participation in AI-supported learning environments.

For language educators, the framework offers a practical approach to task design that positions learners as critics, designers, and advocates rather than passive consumers of AI-generated content. By encouraging learners to analyze, revise, and transform AI outputs through multiple languages and modes of communication, educators can promote learner agency, critical AI literacy, and deeper engagement with language and culture. Such an

approach is particularly relevant in multilingual and international educational contexts where learners bring diverse linguistic, cultural, and experiential resources to the classroom.

The framework also highlights the importance of addressing issues of equity and representation in AI-supported learning. Educators should critically examine how AI-generated content reflects particular cultural perspectives and create opportunities for learners to challenge, expand, and reinterpret these representations. In doing so, AI can become a tool for inclusion rather than the reproduction of existing inequalities.

Future research should examine the implementation of this framework in classroom settings and explore its impact on critical thinking, multilingual participation, and learner engagement. Additional research is needed to investigate how learners at different proficiency levels interact with AI-supported translanguaging and multimodal tasks, as well as how such approaches may support diverse and historically marginalized learner populations.

In positioning AI as a resource for critical inquiry, multilingual expression, and inclusive participation, this framework contributes a theoretically grounded approach to responsible AI integration in language education.

References

Ahmed, A. R. (2025). Navigating the integration of generative artificial intelligence in higher education: Opportunities, challenges, and strategies for fostering ethical learning. *Advances in Biomedical and Health Sciences, 4*(1), 1–2.

Adami, E., Diamantopoulou, S., & Lim, F. V. (2022). Design in Gunther Kress's social semiotics. *London Review of Education, 20*(1), 41. https://doi.org/10.14324/LRE.20.1.41

Blundell, C. N., Mukherjee, M., & Nykvist, S. (2022). *A scoping review of the application of the SAMR model in research.* Computers and Education Open, 3, 100093.

Cheung, K. K. C., Pun, J., Kenneth-Li, W., & Mai, J. (2025). Exploring students' multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools. *Journal of Science Education and Technology, 34*(1), 284–297. https://doi.org/10.1007/s10956-024-10182-0

Cioè-Peña, M. (2022). TrUDL, a path to full inclusion: The intersectional possibilities of translanguaging and Universal Design for Learning. *TESOL Quarterly, 56*(2), 799–812. https://doi.org/10.1002/tesq.074

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology, 55*(6), 2503–2529. https://doi.org/10.1111/bjet.13460

Danesi, M. (2000). *Semiotics in language education* (Vol. 2). Walter de Gruyter. https://doi.org/10.1515/9783110823080

Donley, K. (2024). Teaching with ChatGPT as a linguistically responsive tool for multilingual learners. *Technology in Language Teaching & Learning, 6*(3), 1719–1719. https://doi.org/10.29140/tltl.v6n3.1719

Evmenova, A. S., Borup, J., & Shin, J. K. (2024). Harnessing the power of generative AI to support all learners. *TechTrends, 68*(4), 820–831.

Galante, A. (2020). Pedagogical translanguaging in a multilingual English program in Canada: Student and teacher perspectives of challenges. *System, 92*, 102274. https://doi.org/10.1016/j.system.2020.102274

García, O., & Kleifgen, J. A. (2019). Translanguaging and literacies. *Reading Research Quarterly, 55*(4), 553–571. https://doi.org/10.1002/rrq.286

Kress, G. (2010) Multimodality: *A social semiotic approach to contemporary communication*. London: Routledge.

Pentedura, R. R. (2013). SAMR: Getting to transformation. Retrieved from http:// www.hippasus.com/rrpweblog/archives/2013/04/16/SAMRGettingToTransformation.pdf

Tang, L., & Su, Y. S. (2024). Ethical implications and principles of using artificial intelligence models in the classroom: A systematic literature review. *International Journal of Interactive Multimedia and Artificial Intelligence, 8*, 25. https://doi.org/10.9781/ijimai.2024.02.010

Wanselin, H., Danielsson, K., & Wikman, S. (2022). Analysing multimodal texts in science—A social semiotic perspective. *Journal of Science Education and Technology, 31*, 284–297. https://doi.org/10.1007/s10956-021-09930-6

Bionote

Fati Osuman obtained her PhD in Spanish (Linguistics) from the University of Ottawa, Canada. She is a language educator and researcher with over seven years of experience teaching Spanish and English across diverse levels and settings in Canada, the United States, and Ghana. Her research explores second language education and assessment, particularly where they intersect with inclusive technology and AI. Her dissertation examined student engagement in different Spanish learning formats, and her recent work investigates how generative AI can support language learning through translanguaging and plurilingual approaches.

#24 Activity Theory as a Lens for Examining Teachers' Enactment of the Competency-Based English Teaching Curriculum

Thanh Vinh Chau[1], Thao Quoc Tran[2]

Abstract

Competency-based reform has increasingly influenced language education, placing communicative competence at the center of curriculum development. In Vietnam, this orientation is reflected in the Competency-Based English Teaching Curriculum (CBETC). Despite growing interest in curriculum reform, limited attention has been paid to how CBETC enactment can be understood theoretically. Drawing on curriculum enactment scholarship, which conceptualizes curriculum as something interpreted, negotiated, and reconstructed in practice, this paper contends that existing accounts have tended to emphasize either procedural implementation or teachers' individual beliefs while paying insufficient attention to the wider social and institutional conditions surrounding classroom practice. In response, Activity Theory (AT) is advanced as a framework for examining the enactment of CBETC. Informed by Vygotsky's (1987) sociocultural perspective and Engeström's (2009) third-generation activity theory, the paper conceptualizes curriculum enactment as a socially mediated activity system shaped by the relationships among subject, object, tools, rules, community, and division of labor. The analysis illustrates how tensions and contradictions, particularly those arising from the coexistence of competency-based aspirations and examination-oriented traditions, influence teachers' pedagogical decisions and opportunities for professional action. By bringing AT into dialogue with scholarship on curriculum enactment, the paper offers a more comprehensive account of how curriculum reform is enacted in practice and contributes to ongoing discussions of competency-based English language education. The proposed framework also opens avenues for future empirical investigations of curriculum enactment in diverse educational settings.

Keywords: Activity Theory; curriculum enactment; competency-based education; English language teaching

1. Introduction

1.1. Global shift toward competency-based education

In the context of globalization and the rapid expansion of the knowledge-based economy, general education systems have been witnessing a marked shift from content-oriented models to competency-based approaches. International organizations such as the Organisation for Economic Co-operation and Development (OECD) emphasize that contemporary education should prioritize the development of learners' capacity to apply knowledge in practical situations rather than merely accumulate academic content (OECD, 2018). This global trend has significantly influenced curriculum reforms in many countries, including Vietnam.

Vietnam's 2018 General Education Curriculum represents a pivotal milestone by formally establishing the development of learners' qualities and competencies as its central orientation. Within this framework, English is designed according to communicative competence principles (Ministry of Education and Training [MOET], 2018). The introduction

[1] *Ho Chi Minh City Open University, Vietnam; Coresponding Email: thanhcv.24at@ou.edu.vn*
[2] *Ho Chi Minh City Open University, Vietnam*

of the CBETC, therefore, constitutes not simply an adjustment of content but a fundamental transformation in pedagogical orientation—shifting the focus from the transmission of linguistic knowledge to the development of learners' ability to use language effectively in authentic contexts.

1.2. The complexity of curriculum enactment

However, translating CBETC into classroom practice is not a linear process. Research on curriculum enactment suggests that teachers do not merely "implement" policy documents as written; instead, they interpret, adapt, and reconstruct the curriculum through their everyday instructional practices (Spillane, Reiser, & Reimer, 2002). The notion of enactment thus more accurately captures the dynamic and interactive nature of this process.

Teachers play a central role in transforming competency goals into concrete teaching activities while being influenced by multiple factors, such as professional beliefs, school culture, examination pressures, material conditions, and assessment policies (Priestley, Biesta, & Robinson, 2015). The interweaving of individual and contextual factors renders the enactment of CBETC a complex professional undertaking rather than a straightforward act of policy compliance.

1.3. The need for a comprehensive analytical lens

Linear approaches or those that focus solely on teachers' individual cognition often struggle to capture the systemic and evolving nature of curriculum enactment. A more comprehensive analytical framework is therefore required—one capable of examining the subject, mediating tools, rules, community, and the broader social relations that shape practice. In this regard, Activity Theory (AT), rooted in the socio-cultural tradition of Vygotsky (1987) and further elaborated in its third generation by Engeström (2009), provides such an approach. AT conceptualizes teaching practice as an activity system in which internal contradictions drive change and development.

1.4. Purpose of the paper

Building on the above arguments, this paper aims to examine the theoretical foundations of activity theory, analyze its role as an analytical lens for investigating teachers' enactment of CBETC, and discuss its theoretical and practical implications for curriculum development and teacher professional development.

2. Theoretical Background

2.1. CBETC and the motivation for its implementation

A CBETC can be understood as a curriculum approach in which the goal of instruction extends beyond the transmission of linguistic knowledge to the development of learners' ability to use English in specific communicative situations. From a competency-based perspective, learners are expected to mobilize knowledge, skills, and attitudes to accomplish learning tasks and address practical problems. This approach is consistent with global educational trends, as the OECD emphasizes that contemporary education should prepare learners with the capacity to act, adapt, and participate effectively in future social life (OECD, 2018).

In the Vietnamese context, CBETC is clearly reflected in the 2018 General Education Curriculum for English, issued together with Circular No. 32/2018/TT-BGDĐT. The curriculum states that English contributes to the formation and development of communicative competence through the four skills of listening, speaking, reading, and writing, while also supporting the development of learners' qualities, general competencies, and capacity for

lifelong learning (MOET, 2018). Accordingly, the focus of English language teaching has shifted from memorizing vocabulary, grammar, and test-taking techniques to organizing learning activities that enable students to use language purposefully, contextually, and meaningfully. Recent reviews of CBETC literature in Vietnam further indicate that this shift requires substantial changes in teachers' professional roles, moving from knowledge transmitters to facilitators of communicative learning, learning designers, and assessment mediators (Chau & Tran, 2026). This orientation is also aligned with the spirit of the Common European Framework of Reference for Languages, which emphasizes descriptions of language proficiency based on what learners are able to do in communication (Council of Europe, 2020).

However, CBETC does not automatically translate into classroom practice. Teachers are required to interpret competency-based objectives, select appropriate teaching materials, design communicative tasks, organize classroom interaction, and assess students' progress. Therefore, the enactment of CBETC should be viewed as a process of professional construction shaped by teachers' beliefs, school conditions, assessment pressures, and learning culture.

In contemporary educational research, curriculum is no longer understood as a static document or a fixed body of content to be transmitted directly from policy to classroom practice. Rather, it is increasingly viewed as a social and situated practice constructed, interpreted, and reconstructed through the interaction among teachers, students, and contextual conditions. From the perspective of "curriculum as practice," curriculum comes into being only when it is enacted in specific educational settings (Priestley, Biesta, & Robinson, 2015). In this sense, teachers are not passive implementers of prescribed policy documents; they are professional agents who interpret curricular intentions, select appropriate content, adapt pedagogical strategies, and respond to learners' actual needs (Chau & Tran, 2026).

The enactment of a CBETC should therefore be understood as a dynamic process through which teachers transform competency-oriented goals into concrete classroom practices. This understanding is also reflected in recent studies on CBETC enactment, which identify teachers' beliefs, classroom practices, and contextual conditions as interconnected dimensions shaping curriculum enactment processes (Chau & Tran, 2025b). Unlike traditional content-based approaches that mainly emphasize the transmission of grammar, vocabulary, and linguistic knowledge, CBETC requires teachers to organize learning activities that develop students' ability to use English meaningfully in communicative contexts. This means that teachers must translate broad curricular objectives, such as communicative competence, learner autonomy, collaboration, and problem-solving, into instructional tasks, classroom interaction, assessment practices, and learning support.

The notion of enactment is particularly useful because it moves beyond the linear idea of implementation. While implementation often suggests a direct movement from curriculum policy to classroom application, enactment emphasizes interpretation, negotiation, adaptation, and reconstruction (Spillane, Reiser, & Reimer, 2002). In the context of CBETC, teachers are simultaneously shaped by official curriculum documents and actively involved in re-creating them through their beliefs, pedagogical experience, professional knowledge, school culture, examination pressure, available resources, and students' learning characteristics.

Therefore, the enactment of CBETC is not merely a technical matter of following curriculum guidelines. It is a complex professional activity situated at the intersection of individual agency and broader institutional structures. Teachers must balance competency-oriented expectations with contextual realities, including large class sizes, limited teaching time, traditional assessment practices, and learners' varying levels of English proficiency. To understand this multidimensional process, a theoretical framework is needed that connects

teachers' classroom actions to the social, cultural, and institutional conditions that shape them. Activity Theory provides a particularly relevant foundation for examining the enactment of CBETC as a socially mediated, contextually situated activity system.

2.2. Overview of activity theory

2.2.1. Sociocultural foundations (Vygotskian roots)

Activity Theory originates from the sociocultural tradition of Vygotsky, who argued that higher mental functions are mediated by cultural tools and signs within social contexts (Vygotsky, 1978). Rather than conceptualizing cognition as a purely internal process, Vygotsky maintained that human action is always shaped by culturally constructed means such as language, symbols, documents, and material artifacts. The concept of mediation thus becomes foundational to AT, underscoring the idea that individuals do not interact directly with the world but do so through mediating tools.

Moreover, tool use does not merely facilitate action; it also transforms the subject who acts. Tools are simultaneously products of cultural-historical development and instruments that shape how individuals think and behave. This insight is particularly significant in educational settings, where curricula, textbooks, instructional materials, and assessment policies function as mediating tools within teaching practice.

In addition, learning is conceptualized as a socially situated process. Knowledge is constructed through interaction within communities of practice rather than existing independently within individuals. This perspective provides a basis for analyzing the classroom as a social activity system in which teachers and students collaboratively construct meaning.

2.2.2. First and second generations of activity theory

Building on Vygotsky's foundational principles, Leontiev extended AT by distinguishing between activity, action, and operation, thereby emphasizing the goal-directed nature of human activity (Leontiev, 1978). This distinction suggests that activity is always oriented toward an object, an element that generates both motivation and meaning for human action.

Subsequently, on this theoretical foundation, Yrjö Engeström developed the second generation of AT by constructing the well-known expanded triangular model of the activity system (Engeström, 1987). This model comprises the following core components:

- Subject
- Object
- Tools (mediating artifacts)
- Rules
- Community
- Division of labor

By incorporating social dimensions such as rules, community, and division of labor, AT moves beyond individual, level analysis and enables the examination of action in relation to institutional and cultural structures. In educational research, this model is particularly valuable for analyzing teaching practice as a complex system in which teachers do not act independently but are influenced by multiple interacting forces.

2.2.3. Engeström's third-generation activity theory

In the third generation, Engeström (2009) further expanded AT to analyze interactions among multiple activity systems. Rather than focusing on a single activity system, as in earlier versions, Engeström proposed a model of interacting activity systems, which is especially suitable for contexts of educational reform, where policy, school, and classroom can be conceptualized as interconnected systems.

In addition, a significant contribution of the third generation is the concept of expansive learning. According to this perspective, learning is not merely the acquisition of existing knowledge but a process of reconstructing the object of activity through the resolution of internal contradictions. Contradictions are viewed as the driving force for change and development within activity systems. For example, in the context of the CBETC, the contradiction between the goal of developing communicative competence and grammar-oriented assessment practices may create tension while simultaneously opening up opportunities for innovation in teaching practice.

By emphasizing the role of contradictions and transformation, third-generation AT provides a dynamic analytical lens through which curriculum enactment can be understood not as passive adaptation but as a developmental process of reconstruction. This perspective is particularly relevant to research on the enactment of CBETC, where teachers continuously adjust their practices to respond to policy demands and contextual constraints.

In sum, Activity Theory offers a comprehensive theoretical framework for analyzing teachers' curriculum enactment as a socially situated activity system, highlighting the roles of mediation, context, and contradiction in professional development and educational change.

2.3. Applying activity theory to teachers' enactment of CBETC

2.3.1. Conceptualizing CBETC enactment as an activity system

Based on Engeström's (2009) analytical framework, the process of teachers enacting the CBETC can be conceptualized as an activity system in which its components interact dynamically.

First, the subject of the system is English language teachers—the professional agents who directly interpret and actualize the curriculum in classroom practice. They do not merely transmit content; instead, they make pedagogical decisions grounded in their experience, beliefs, and contextual conditions. Second, the object of the activity is the development of students' English communicative competence in alignment with the orientation of the new general education curriculum. This object is dynamic and may be reconstructed in practice, for instance, when teachers shift their focus from communicative competence development to ensuring examination performance. Third, the tools include textbooks, curriculum guidelines, policy documents, instructional technologies, and pedagogical strategies. From Vygotsky's (1978) perspective, these tools function as mediating artifacts that shape how teachers organize learning activities.

In addition, rules encompass assessment regulations, school culture, professional standards, and societal expectations. These rules either support or constrain the enactment of competency-oriented activities. The community consists of students, colleagues, school leaders, and parents—stakeholders who participate in or influence the enactment process. Finally, the division of labor reflects the allocation of roles among teachers, subject departments, and school leadership in organizing and supervising instructional activities. Conceptualizing CBETC enactment as an activity system allows for the simultaneous analysis

of individual and structural factors, rather than reducing the issue to teachers' personal competence alone.

2.3.2. Mediations and interactions within the activity system

Within this system, components do not exist in isolation but continuously interact and mediate one another. Instructional tools, such as competency-oriented textbooks, may encourage teachers to organize communicative activities, project-based learning, or formative assessment. However, tool use is not mechanical; teachers may reinterpret materials in light of their own experience and beliefs. According to Spillane et al. (2002), policy implementation is inherently a process of "sense-making," in which teachers interpret and adapt policy messages within specific contexts.

Rules also function as critical mediating elements. In educational contexts strongly influenced by high-stakes examinations, the pressure to achieve high scores on standardized tests may lead teachers to prioritize grammatical structures over the development of communicative competence (Chau & Tran, 2025a). This illustrates how the interaction between the object (competency development) and rules (knowledge-oriented assessment) can transform the orientation of activity.

Moreover, community and division of labor shape how teachers enact the curriculum. Professional support from subject departments, structured professional meetings, and leadership direction may create enabling conditions for methodological innovation. Conversely, a lack of coordination or unclear role allocation may undermine the effectiveness of enactment. As argued by Priestley et al. (2015), teacher agency is always formed within an ecological relationship between individuals and their contexts.

2.3.3. Contradictions and tensions in CBETC enactment

One of the key contributions of AT is its conceptualization of contradictions as drivers of change (Engeström, 2009). In the context of CBETC, the following three typical contradictions can be identified.

The first concerns the tension between the goal of communicative competence development and an assessment system oriented toward grammatical knowledge. Within the AT framework, this represents a contradiction between the object and rules of the activity system. In practice, when examinations continue to emphasize linguistic structures, teachers may adjust their instructional activities to meet assessment demands, thereby weakening the coherence of the competency-based curriculum (Chau & Tran, 2025a).

The second contradiction lies between innovative pedagogical approaches (e.g., project-based learning, group work, formative assessment) and contextual constraints such as large class sizes, limited instructional time, or insufficient technological resources. This reflects tensions among tools, rules, and the material conditions embedded within the community.

The third is an internal contradiction among teachers, specifically between beliefs in communicative language teaching and entrenched traditional teaching habits. According to Leontiev (1978), the motive of activity may shift when the object is reconstructed; thus, the tension between long-term professional goals and short-term performance pressures may lead to adjustments in pedagogical action. These contradictions should not be viewed solely as barriers but as potential sources for restructuring the activity system.

2.3.4. Expansive possibilities and transformation

From the perspective of expansive learning proposed by Engeström (2009), when contradictions are identified and critically reflected upon, they may become resources for innovation. For example, pressure arising from misalignment between competency goals and assessment formats may motivate teachers to seek alternative assessment approaches, such as task-based or portfolio assessment. However, such transformation requires specific enabling conditions, including: professional spaces for teachers to discuss and reflect; more flexible assessment policies; school leadership that encourages pedagogical experimentation; and the development of professional learning communities.

Accordingly, only when the components of the activity system are more harmoniously aligned, for instance, when rules are consistent with the object and tools genuinely support competency goals, CBETC enactment achieves greater sustainability. Therefore, Activity Theory not only facilitates the description of practice but also provides an analytical foundation for guiding systemic change.

In sum, applying Activity Theory enables the conceptualization of teachers' enactment of CBETC as a complex activity system in which mediating factors, contradictions, and social interactions jointly shape instructional practice. This approach clarifies the dynamic, contextualized, and developmental nature of curriculum enactment within the broader context of educational reform.

2.4. Roles of Activity Theory in Understanding CBETC Enactment

2.4.1. Providing a holistic analytical framework

One of the significant contributions of AT to research on CBETC enactment lies in its capacity to provide a comprehensive analytical framework. Rather than focusing solely on teachers' cognition or individual beliefs, AT situates teaching practice within a broader activity system, where the subject's actions are always mediated by tools, rules, community, and the division of labour (Engeström, 1987).

In many studies on curriculum reform, cognitively oriented approaches tend to explain variations in practice primarily through "teacher beliefs" (Borg, 2003; Fives & Gill, 2015; Spillane, Reiser, & Reimer, 2002). However, as Spillane, Reiser, and Reimer (2002) argue, policy implementation is not merely a matter of understanding or misunderstanding policy texts; it is a process of interpretation situated within specific social contexts. AT enables researchers to move beyond this reductionist view by conceptualizing teachers as acting subjects embedded in institutional and cultural ecosystems.

When applied to CBETC, Engeström's (1987) analytical framework facilitates the simultaneous integration of structural factors—such as assessment regulations, professional standards, and examination pressures—with cultural elements, including established teaching traditions, parental expectations, and students' learning habits. Consequently, enactment is not reduced to a question of the degree of compliance with curricular prescriptions; rather, it is understoodas a socially negotiated and continuously restructured practice.

This holistic approach is particularly well-suited to the context of educational reform in Vietnam, where CBETC is implemented within a system long shaped by an examination-oriented model of teaching and learning. AT makes it possible to analyse the relationship between competency-oriented goals and existing assessment structures as components of a single activity system, rather than as isolated factors.

2.4.2. Capturing complexity and non-linearity

A second contribution of AT lies in its capacity to capture the complexity and non-linearity of curriculum enactment. In practice, enactment does not unfold according to a simple cause–and–effect sequence. Instead, it is shaped through dynamic interactions among multiple components within the activity system.

Grounded in Vygotsky's (1978) sociocultural theory, AT posits that human action is always mediated by tools and social context. This implies that changes in one component (e.g., modifications in assessment formats) may lead to adjustments in other components (e.g., instructional strategies). The expanded triangular structure of the activity system provides a means of representing these multidirectional interactions.

Moreover, third-generation AT emphasizes interactions among multiple activity systems (Engeström, 2009). In the case of CBETC, at least three interrelated systems can be identified: the policy system (ministerial level), the school system, and the classroom system. These systems do not operate independently; instead, they continuously influence one another. For example, changes in national assessment policy may reshape teachers' objectives of activity at theclassroom level.

AT thus helps avoid a linear conception in which policy is "transmitted" from the top and implemented intact at lower levels. Instead, enactment is understood as the outcome of ongoing negotiations between structural forces and individual agency. This perspective is critical in the study of CBETC, a reform that requires coordinated changes in objectives, content, pedagogy, and assessment.

2.4.3. Identifying sources of change and development

Another notable contribution of AT is its ability to identify sources of change and development through the concept of contradictions. According to Engeström (2009), internal contradictions within an activity system are not merely obstacles but also drivers of expansive learning.

In the context of CBETC, contradictions—such as the misalignment between competency-oriented goals and knowledge-focused testing practices—can serve as productive points of analysis. Rather than interpreting such tensions as evidence of teachers' failure, AT encourages their recognition as indicators of systemic imbalance. Analysing contradictions between the object (competency development) and the rules (traditional assessment practices) makes it possible to illuminate structural constraints that must be addressed for the curriculum to function effectively.

This perspective carries important implications for teacher professional development. When teachers engage in reflective examination of contradictions in their practice—for instance, through professional learning communities or lesson study—they may reconstruct the object of their activity and develop new forms of practice. Professional growth, therefore, is not simply the accumulation of instructional techniques; it is a transformative process that reconfigures the activity system in alignment with curricular goals.

AT thus offers an analytical lens that is both diagnostic and developmental. It enables the identification of bottlenecks in CBETC enactment while also pointing toward conditions for sustainable change, including adjustments to assessment regulations, strengthened professional support, and the cultivation of collaborative learning communities within schools.

In summery, activity theory functions as a robust theoretical framework for understanding CBETC enactment in three respects: it provides a holistic perspective that moves beyond individual cognition; it captures the complexity and non-linearity of the enactment

process; and it identifies sources of change through the analysis of systemic contradictions. In doing so, AT not only clarifies the nature of teaching practice in reform contexts but also informs strategies for curriculum development and teacher professional learning.

3. Methodology

This study was conducted as a theoretical analysis to clarify the role of Activity Theory in examining teachers' enactment of the competency-based English teaching curriculum. Instead of collecting empirical data from classroom settings, the paper focuses on analyzing, synthesizing, and interpreting research studies and theoretical papers related to competency-based curriculum, curriculum enactment, communicative competence in English language teaching, and Activity Theory. This approach is appropriate to the paper's purpose because the main focus of the study is not to measure the extent of curriculum implementation but to develop an analytical perspective capable of explaining the complex, socially situated, and context-dependent nature of CBETC enactment.

The main sources selected for analysis include foundational works on Vygotsky's sociocultural theory and Engeström's activity system model, studies on curriculum enactment by Spillane and colleagues, and documents related to Vietnam's 2018 General Education Curriculum for English. These materials were read and analyzed with a focus on identifying key concepts, including subject, object, tools, rules, community, division of labor, contradiction, and expansive learning. On this basis, the paper examines how these components can be applied to understand English teachers' instructional activity in the context of the shift from content-oriented to competency-based teaching.

The analytical procedure was carried out in three main steps. First, the paper clarifies the conceptual foundations of CBETC and the issue of curriculum enactment. Second, the components of Activity Theory are systematized and connected to the context of English language teaching in Vietnamese high schools. Third, the paper analyzes typical contradictions in the enactment of CBETC, particularly the misalignment between the goal of developing communicative competence and the pressure to test and examine linguistic knowledge. In doing so, the study proposes Activity Theory as an analytical lens that connects teachers' individual actions to broader institutional, cultural, and social conditions.

4. Results

4.1. Implications

4.1.1. Theoretical implications

The application of AT to the study of teachers' enactment of CBETC generates significant theoretical implications for the field of curriculum studies. First, AT reinforces the view of curriculum enactment as a structured social activity rather than a series of discrete individual actions. Drawing on Yrjö Engeström's (2009) model of the activity system, curriculum research can simultaneously analyse the subject, object, tools, rules, and community as elements of an integrated whole. This perspective broadens traditional approaches that have tended to focus primarily on content analysis or teachers' beliefs.

Second, AT contributes to research on competency-based education by offering a lens through which to interpret the discrepancy between policy intentions and classroom practice. In reform contexts, the shift from a knowledge-based model to a competency-based model frequently encounters an "implementation gap." Through the concepts of contradictions and expansive learning, AT enables the identification of internal systemic tensions as drivers of development rather than merely as indicators of policy failure (Engeström, 2009).

Furthermore, Vygotsky's (1978) sociocultural foundation underscores the role of mediation and context in the construction of practice. This reinforces the argument that research on CBETC must move beyond a purely technical perspective and instead examine the relationship between pedagogical action and institutional structures. In this sense, AT functions not only as a descriptive tool but also as a theoretical foundation that orients curriculum studies toward a systemic and developmental approach.

4.1.2. Practical implications

In addition to its theoretical contributions, the application of AT yields several practical implications for stakeholders involved in CBETC implementation.

For curriculum developers, AT suggests that curriculum design cannot be separated from the conditions of enactment. Consistency between competency-oriented goals (object) and the assessment system (rules) is essential to minimizing contradictions within the activity system (Chau & Tran, 2025a). This requires coherent alignment among content, pedagogy, and assessment, rather than reform limited to policy documents alone.

For school leaders, AT highlights the significance of community and division of labour in supporting teachers. Establishing professional learning communities, creating spaces for reflection, and encouraging pedagogical experimentation can assist teachers in addressing contradictions in practice. As Priestley, Biesta, and Robinson (2015) argue, teacher agency is formed within an ecological relationship between individuals and their contexts; consequently, the organizational environment plays a decisive role in the success of CBETC enactment.

Regarding teacher professional development, AT advocates a system-oriented, reflective approach rather than merely providing instructional techniques. By analysing contradictions between curricular goals and classroom realities, teachers can engage in expansive learning, reconstruct the object of their activity, and adjust pedagogical strategies. Professional development thus becomes a process of transforming the activity system, rather than simply enhancing individual skills.

4.1.3. Directions for future research

Although AT provides a robust theoretical foundation, further empirical validation is needed in specific contexts. The use of mixed-methods designs or in-depth case studies may clarify how components of the activity system interact and evolve over time. In addition, the development and validation of questionnaire instruments could provide a quantitative basis for examining factors affecting teachers' enactment of CBETC (Chau & Tran, 2025b). This is particularly important in the Vietnamese context, where disparities in infrastructure, learning culture, and local policy conditions remain substantial. Moreover, integrating AT with other theoretical frameworks—such as teacher agency or ecological approaches, may enrich the analysis and deepen understanding of teachers' roles within reform systems. Future empirical studies may also employ validated questionnaires to measure the extent to which beliefs, practices, impetus, and hindrances influence CBETC enactment (Chau & Tran, 2025b).

5. Discussion

A further contribution of this paper lies in highlighting the analytical value of contradictions in understanding CBETC enactment. Within Engeström's (2009) framework, contradictions are viewed as historically accumulated tensions that emerge among elements of an activity system and create possibilities for transformation. In the Vietnamese context, one of the most visible tensions concerns the coexistence of competency-oriented curricular goals

and examination practices that continue to reward the reproduction of linguistic knowledge. Such tensions suggest that difficulties in curriculum enactment cannot be reduced to teachers' individual limitations or resistance to change. Instead, they arise from the interaction between policy aspirations and institutional conditions.

This interpretation is consistent with a growing body of research in language education that has applied Activity Theory to investigate educational innovation and change. For example, Blin and Appel (2011) demonstrated that tensions among technological tools, institutional expectations, and established pedagogical routines significantly influence how language teachers respond to innovation. Their work illustrates how pedagogical change is shaped not only by teachers' intentions but also by the relationships among multiple elements within an activity system. Viewed from this perspective, the challenges associated with CBETC enactment reflect broader patterns identified in Activity Theory research, where curriculum innovation is frequently mediated by contextual constraints, competing priorities, and institutional arrangements.

The discussion also suggests that Activity Theory offers a broader explanatory framework than approaches centered primarily on teacher beliefs or cognition. While studies such as Borg (2003) and Fives and Gill (2015) have generated valuable insights into teachers' thinking, AT highlights the social, cultural, and organizational conditions in which those beliefs are formed and enacted. Rather than treating beliefs as isolated determinants of practice, Activity Theory positions them within a network of mediating influences, including institutional regulations, professional communities, material resources, and assessment systems. This perspective allows curriculum enactment to be understood as a socially situated process rather than solely an individual cognitive endeavor.

Consequently, supporting the enactment of CBETC requires more than methodological training. Sustainable curriculum reform depends on stronger coherence among curriculum objectives, assessment systems, school leadership, professional collaboration, and classroom realities. This view is also compatible with Priestley, Biesta, and Robinson's (2015) ecological understanding of teacher agency, as well as Chau and Tran's (2025b, 2026) argument that curriculum enactment should be examined through the relationship between teachers and the broader environments in which they work. Taken together, these perspectives reinforce the argument that teachers should be viewed as active participants in interconnected activity systems whose practices evolve through ongoing engagement with contextual opportunities and constraints. Effective enactment of CBETC, therefore, depends not only on teachers' professional capacities but also on the alignment of the wider conditions that support curriculum reform.

6. Conclusion

This article has argued that Activity Theory provides a productive analytical lens for examining teachers' enactment of CBETC. Through the activity system model, AT conceptualizes enactment as a complex social practice in which elements such as subject, tools, rules, and community continuously interact.

This systemic perspective moves beyond linear views of implementation by emphasizing the dynamic, non-linear, and contextually situated nature of teaching practice. At the same time, the concepts of contradictions and expansive learning open possibilities for identifying and guiding change within reform processes.

In the context of CBETC implementation, careful analysis of contextual factors—from assessment policy and school culture to classroom conditions—is a prerequisite for achieving

competency-development goals. The application of activity theory, therefore, not only facilitates a deeper understanding of practice but also informs strategies for sustainable improvement in general education.

Acknowledgments

The authors would like to express their sincere gratitude to Ho Chi Minh City Open University for providing academic support and a research environment that enabled the development of this paper. The authors are also grateful to colleagues and scholars in the field of English language education whose discussions and academic insights helped shape the theoretical orientation of this study. Special thanks are extended to those whose work on competency-based education, curriculum enactment, and Activity Theory provided an important foundation for the arguments presented in this article. Any remaining limitations or shortcomings are solely the responsibility of the authors.

Reference

Borg, M. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81–109. https://doi.org/10.1017/S0261444803001903

Blin, F., & Appel, C. (2011). Computer supported collaborative writing in practice: An activity theoretical study. *CALICO Journal, 28*(2), 473–497. https://doi.org/10.11139/cj.28.2.473-497

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. https://www.coe.int/en/web/common-european-framework-reference-languages

Chau, V. T., & Tran, Q. T. (2025a). Assessment and evaluation in competency-based English teaching: Affordance and challenges. In *Proceedings of the 13th OpenTESOL International Conference 2025: Embracing voices: Diversity, agency and growth in language education* (pp. 661–674). Ho Chi Minh City Publishing House of Economics. https://opentesol.org/past-conferences/?utm_source=chatgpt.com

Chau, V. T., & Tran, Q. T. (2025b). The questionnaire about factors affecting enactment of competency-based English teaching curriculum: A developing and validating method. *International Journal of TESOL Studies, 250802*, 1–25. https://doi.org/10.58304/ijts.250802

Chau, V. T., & Tran, Q. T. (2026). Revisiting Teacher's Roles under Competency-Based English Teaching Curriculum: A Systematic Literature Review. *VNU Journal of Science: Education Research*, *42*(1), 24-40. https://doi.org/10.25073/2588-1159/vnuer.5364.

Engeström, Y. (2009). Expansive learning: Toward an activity-theoretical reconceptualization. *Journal of Education and Work, 14*(1), 133–156. https://doi.org/10.1080/13639080020028747

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit. https://lchc.ucsd.edu/mca/Paper/Engestrom/Learning-by-Expanding.pdf

Fives, H., & Gill, M. G. (2015). *International handbook of research on teachers' beliefs*. New York, NY: Routledge. https://www.researchgate.net/profile/Jeppe-Skott/publication/266774641_The_promises_problems_and_prospects_of_research_on_teachers'_beliefs/links/56f2366508ae1cb29a3d2543/The-promises-problems-and-prospects-of-research-on-teachers-beliefs.pdf

Ministry of Education and Training (MOET). (2018). *General education curriculum – English (attached to Circular No. 32/2018/TT-BGDĐT)*. MOET. https://thbausen.hcm.edu.vn/chuong-trinh-chinh-thuc/chuong-trinh-giao-duc-pho-thong-2018-mon-tieng-anh/ctmb/85329/442876

OECD. (2018). *The future of education and skills 2030*. https://www.oecd.org/education/2030-project/

Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury. https://www.bloomsbury.com/uk/teacher-agency-9781472534668/

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431. https://sci-hub.vn/10.3102/00346543072003387

Vygotsky, L. S. (1978). *Mind in society: The Development of Higher Psychological Processes*. Harvard University Press. https://home.fau.edu/musgrove/web/vygotsky1978.pdf

Bionote

Mr. Chau Vinh THANH is an English lecturer pursuing a PhD in TESOL at Ho Chi Minh City Open University. His main research interests are English learning and teaching, curriculum design, teaching methods, autonomous learning and teaching. ORCID: https://orcid.org/0009-0001-1508-42811

Associate Professor ***Dr. Tran Quoc Thao*** is currently working at the Faculty of Foreign Languages, Ho Chi Minh City Open University. He has been teaching English for 20 years, and he is experienced in teaching English to both local and international students. His research areas are discourse analysis, English language skills, teacher professional development and instructional design model.

#26 Enhancing Coherence and Cohesion in IELTS Writing through Generative AI Feedback: An Exploratory Study

Nhu Hoang Ngoc Quynh[1] & Sa Tran Si Nguyen[2]

Abstract

Improving coherence and cohesion (CC) in academic writing, especially for IELTS preparation, is challenging. It requires constructive feedback and sufficient practice, but time constraints often prevent teachers from properly assessing student's work. Generative AI models like ChatGPT can be an effective supplementary tool to assist students. This study examines how 20 students at Industrial University of Ho Chi Minh leverage this tool to enhance CC in their writing. Data was collected through comparing writing samples before and after ChatGPT intervention. Both versions of the essays were evaluated to identify any positive changes in cohesive devices and idea organization. Furthermore, questionnaires were conducted to gather students' experiences, strategies for using AI-generated feedback. The findings reveal both advantages and challenges of AI-generated suggestions on student's Cohesion and Coherence. This research seeks to provide meaningful insights into the role of generative AI, offering implications for English language teaching and learning.

***Keywords*:** cohesion and coherence, generative AI, ChatGPT, IELTS writing

1. Introduction

In the context of educational globalization and internationalization, IELTS certificate holds a more important role in evaluating a learner's English competency serving educational and professional purposes. Among four skills of the IELTS test, writing tends to be considered as the most difficult skill for EFL learners to acquire and master. According to IELTS performance statistics for Vietnamese test takers in the period 2024 - 2025, the average band score in Listening was 6.5, while Speaking had a lower mean of 5.6. Reading demonstrated relatively strong performance with an average of 6.4, while Writing fell behind at 6.1. This performance gap shows that writing continues to pose a significant challenge for Vietnamese EFL learners, thus emphasising the importance of providing learners with more specialised support and teaching methods aimed at enhancing learners' writing competence (Test taker performance 2024-2025, n.d.). One of the main causes of this difficulty is that writing requires managing both cognitive processes and advanced language use at the same time (Kormos, 2023). In addition to grammar, writing is about control over vocabulary, syntax, and discourse organisation. Learners are expected to effectively generate, organise, and articulate ideas in a logical sequence. For this reason, coherence and cohesion criteria, which refers to connections and organisation of ideas (Halliday & Hasan, 2014), are considered crucial and accounts for 25% of IELTS's total writing score. Hyland (2019) also affirms that in academic contexts like the IELTS Writing Task 2, coherence and cohesion are the primary evaluation (Hyland, 2019). Despite their importance, coherence and cohesion present considerable challenges for many EFL learners, particularly at intermediate and upper-intermediate levels (Banerjee & Wall, 2006; Schoonen et al., 2003). Traditional writing assessment approaches often focus on grammar and vocabulary development but fail to sufficiently provide individualised feedback on higher order writing skills such as organisation and textual flow (Hyland, 2003; Bitchener

[1] *Lecturer, Faculty of Foreign Languages, Industrial University of Ho Chi Minh, Ho Chi Minh City, Vietnam; Email: hnqnhu188@gmail.com*
[2] *Lecturer, Foreign Language Center, Ho Chi Minh University of Technology, Ho Chi Minh City, Vietnam*

& Ferris, 2012). In many instructional settings, teachers are limited by large-sized classes, limited time, and heavy focus on test-taking strategies, leaving little room for detailed feedback on discourse-level features (Bitchener & Ferris, 2012). As a result, learners frequently struggle to apply their linguistic knowledge into well-structured academic essays. Moreover, the subjective and complex nature of coherence and cohesion makes them difficult to address with traditional corrective feedback methods. Learners are usually given general comments such as "improve organisation" or "use more linking words" which do not provide detailed and practical guidance for revision. Consequently, there are limited opportunities for students to engage in cycles of writing and reflection, which are crucial for writing development (Hyland & Hyland, 2006).

Generative artificial intelligence (AI), particularly tools such as ChatGPT or Gemini, can serve as a promising solution. These AI models can produce contextual and human-like feedback in terms of organisation, argumentation, and clarity (Swiecki et al., 2022; Banihashem et al., 2024.; Shi & Aryadoust, 2024). When providing learners with immediate and personalised feedback, AI tools may create new opportunities for improving coherence and cohesion in IELTS writing, while also supporting learner autonomy and reflection. As such, the integration of GenAI in learning and teaching writing skill has increasingly gained attention from scholars in recent years.

Even though generative AI tools such as ChatGPT raise new possibilities, most studies focus solely on aspects such as accuracy in grammatical structures, profound vocabulary range, or overall quality of essays, ignoring discourse-level features, particularly cohesion and coherence. Furthermore, there is still limited empirical research on their effectiveness in improving higher order writing skills in IELTS preparation contexts. This study bridges this gap by investigating how AI-generated feedback influences learners' ability to enhance coherence and cohesion in IELTS Writing Task 2 essays and how learners perceive this type of support. In particular, the study seeks to explore how AI-generated responses can function as a tool to guide learners in better organising their ideas and ensuring greater clarity in their arguments. Beyond textual outcomes, the study also aims to explore learners' perceptions of AI feedback in terms of its usefulness and reliability as well as potential limitations (Steiss et al., 2024). Therefore, it contributes a balance between empirical evidence of writing improvement and an understanding of learners' learning experiences.

To achieve these objectives, the study addresses two central research questions:

1. To what extent does generative AI feedback enhance coherence and cohesion in IELTS Writing Task 2 essays?
2. How do EFL learners perceive the usefulness and reliability of generative AI feedback for improving their IELTS writing performance?

2. Literature Review

2.1. Feedback Intervention Theory (FIT)

Feedback has always been a pivotal aspect to enhance study quality, especially in learning and teaching writing. Of all theories regarding the role of feedback, Feedback Intervention Theory (FIT), proposed by Kluger and DeNisi (1996), posits that effectiveness of feedback is maximised when it helps learners realize the gap between their current performance and desired goals, thereby opting for plausible reactions to bridge this gap. According to FIT, feedback does not always work, if vague or overly critical ones are given, it causes

performance to decline. Instead, specific and corrective mistakes help learners know exactly how to fix the problem, hence improving the way learners perform.

In academic contexts, feedback is especially important since learners hardly realize weaknesses related to coherence and cohesion. As such, feedback that is specific, timely and able to orient reactions allows learners to make necessary adjustments in editing essays. The blooming growth of Generative Artificial Intelligence (GenAI) offers a new feedback form meeting various features of effective feedback following FIT. Unlike the traditional automatic writing assessment system which mainly focuses on grammar and vocabulary mistakes, large language models such as ChatGPT are capable of analyzing how arguments are structured, pointing out problems in flow of ideas and suggest how to improve them. Thanks to the ability of giving immediate and individualized feedback, GenAI is expected to assist learners in improving discourse-level aspects, especially coherence and cohesion. However, FIT also emphasizes that the efficacy of feedback depends not only on the quality of feedback but also on how users make use of feedback. As a result, to evaluate pedagogical value of GenAI, it is essential to consider both the improvement in writing products and the way learners perceive this feedback tool.

2.2. Coherence and Cohesion

In academic writing, beside concepts that are regularly emphasized by educators such as grammatical range, lexical resource, task response, and accuracy, the aspects of coherence and cohesion hold paramount significance as well. These two elements determine whether an essay is logical and whether ideas flow smoothly. It is no exaggeration to state that coherence and cohesion are the cornerstone of a high-band essay. Therefore, sufficient practice in organizing ideas logically and employing appropriate lexical and grammatical devices to enhance connection of sentences in text is indispensable.

Halliday and Hasan (2014) put forward intertwined terms cohesion and coherence. Cohesion concerns semantic meanings of a text. There are two major offshoots: lexical cohesive devices (reiteration and collocation) and grammatical cohesive devices (reference, substitution, ellipsis and conjunctions) to connect ideas both sentence level and paragraph level. Through cohesion texts are linguistically linked, forming a cohesive and interdependent structure (Janebi Enayat, 2025).

On the other hand, coherence refers to the unity of ideas in the paragraph. In this regard, the paragraph is linked smoothly and logically from topic sentence, to supporting sentence, detail sentences and concluding sentences. Oshima and Hogue (2006) propose four ways to achieve coherence: repeating key nouns, maintaining a consistent tone, using transition signals to connect ideas, and arranging ideas logically. When coherence is established, readers can effortlessly grasp the intended meaning of the text.

In the context of IELTS writing, achieving a higher band score (at least Band 6) means students have a good command of using Coherence and Cohesion as the band descriptor requires students to present their ideas in a clear, logical, and coherent manner. Javahery (2025)'s research put forward that essays which fail to exhibit clear progression of ideas and varied cohesive devices are likely to be given Band 6, instead of Band 7 or above. However, many students tend to focus primarily on other aspects of writing and overlook coherence and cohesion, which ultimately reduces their overall performance. There are innumerable research confirming the struggle of most students with utilizing cohesion and coherence in academic writing. A study of Rahman et al., (2023) unveil that although undergraduates majoring in English at Islamia College Peshawar attempt to use cohesive devices in their academic writing,

their application is mostly inappropriate, hindering achieving plausible coherence. This claim is echoed by Trang et al. (2023) and Anindita (2024), in which students are reported to encounter problems such as overuse and misuse conjunctions, reference, repetition and synonyms, and transition signals while underuse cohesive devices like substitution, and ellipsis as well as other types of coherent devices. Furthermore, some students admit they have yet developed a solid grasp of cohesion and coherence, with a few even confusing the two terms (Diep & Diep ,2024). The struggle in employing cohesion and coherence can be attributed to three main reasons: lack of motivation, limited lecture duration and prior English knowledge (Leli, 2020).

Many efforts have been devoted to improving cohesion and coherence in students' academic essays. Previous research has examined whether explicit and implicit instruction in cohesion and coherence can improve students' academic writing. Li and Chen (2026)'s research, which employs a controlled experimental design with 64 upper intermediate-level undergraduate students at a Chinese university, figures out the noticeable improvement in students' academic writing skills when coherence is explicitly taught. Similarly, writing abilities of Iranian EFL learners are proven to be enhanced when they are exposed to explicit instruction (Behbahani et al., 2018). Bahang et al. (2021) also found that teachers can raise students' awareness of how to structure essays logically by explicitly introducing effective coherence-building methods such as theme-rheme progression, which ultimately gives rise to higher writing performance. However, time constraints in classroom settings remain a persistent challenge, as teachers are often unable to provide comprehensive feedback on every essay or give tailored guidance on logical organization to each individual student's writing.

To address this limitation, alternative supports such as automated writing evaluation (AWE) systems and peer feedback have been proposed. Chen and Cui (2022) report that these approaches can foster certain improvements in students' cohesion and coherence. Similarly, Anindita (2024) demonstrates that tools like the Tool for the Automatic Analysis of Cohesion (TAACO), combined with peer feedback, can help identify common cohesive errors and provide systematic evaluations of student writing. Nevertheless, feedback generated by such tools is often too general, insufficiently constructive, and lacking in concrete suggestions for improvement. Peer feedback, while valuable, also has limitations. Termjai (2024) observes that students tend to focus more on grammar and vocabulary than on coherence and unity when providing comments, while Sun et al. (2023) further notes that many learners lack the knowledge necessary to give meaningful corrections or suggestions.

2.3. The integration of ChatGPT in improvement coherence and cohesion

To address these issues, ChatGPT has recently emerged as a viable substitute that is accessible to all students at any time and is increasingly perceived as accurate and reliable. Unlike earlier automated tools, ChatGPT, continuously updated and improved, can generate contextually relevant explanations, suggest personalized revisions, and respond interactively to students' prompts until they are satisfied with their refined drafts (Kohnke, Moorhouse, & Zou, 2023). A myriad of studies have shed fascinating light on the utilization of generative AI, particularly ChatGPT, to enhance learners' writing skills, not only for grammatical or vocabulary correction, but also for cohesion and coherence improvement. Li et al.'s (2023) findings demonstrate that ChatGPT is capable of generating essays that are not only flawless in grammatical structures and lexical resources but also meet academic standards of coherence and cohesion. Herbold et al. (2023) and Tengler and Brandhofer (2025) similarly argue that ChatGPT's essays resemble human writing because they employ logical connections, appropriate transitions, and ensure a smooth flow of ideas between sentences. Furthermore,

ChatGPT's feedback mechanism on coherence and cohesion offers considerable benefits for learners. As Sekewael and Anaktototy (2024) observe, students can input their own drafts and receive personalized, relevant, and real-time feedback. This immediacy enables them to revise their work, learn to think more logically, and strengthen the flow of ideas without having to wait days for teacher feedback, or in some cases, without receiving any feedback at all.

It can be concluded from the review that, although there is growing interest in employing AI-assisted writing tools, most existing research has concentrated on improving surface-level aspects such as grammar, vocabulary, or overall essay structure. Few studies have placed emphasis on the practical application of generative AI in enhancing discourse-level features, particularly cohesion and coherence, which are critical for academic writing and IELTS performance. Moreover, limited attention has been given to learners' real-time interactions with AI tools and their perceptions of such feedback. This highlights the need for research that explores the transformative power of generative AI feedback, specifically ChatGPT, in improving cohesion and coherence in IELTS writing from the learner's perspective.

3. Methodology

3.1 Pedagogical Setting & Participants

Participants were 20 third-year EFL students majoring in English Studies at the Industrial University of Ho Chi Minh City. Each of the participants are expected to reach at least a B1 Intermediate level on the Common European Framework of Reference for Languages (CEFR) since they are in their third year. The research setting was a six-week IELTS Writing Task 2 workshop integrated into the Discourse Analysis course. All students were required to achieve a minimum IELTS score of 6.5 for graduation, which provided a strong motivation for participation in this study.

3.2 Design of the Study

Our research employed a mixed-methods design, following Creswell and Plano Clark (2017). This approach combined quantitative and qualitative methods to investigate the role of generative AI feedback in enhancing the coherence and cohesion in IELTS Writing Task 2 essays. The rationale behind this design lies in the complementary strengths of both data types: quantitative analysis can be used to measure learners' performance improvement, while qualitative inquiry provides deeper insights regarding learners' perceptions and experiences. Following Li et al. (2023), mixed-methods designs are particularly appropriate for applied linguistics research because they allow researchers to reveal the complexity of second language learning phenomena by integrating measurable outcomes with learners' reflective perspectives.

3.3 Data Collection & Analysis

Two distinct IELTS Writing Task 2 essays were used as pre-test and post-test to measure improvement in 20 participants' writing performance. Every essay was evaluated using the official IELTS Writing Band Descriptors with a great emphasis on coherence and cohesion. Participants were provided with automated feedback generated by ChatGPT. The data collected included essay scores and responses from a post-intervention questionnaire designed to record learners' perceptions of AI-mediated feedback in academic writing.

The study lasted six weeks and followed a structured intervention process. During the first week, participants were required to write a pre-test IELTS Writing Task 2 essay, which was assessed using the official IELTS descriptors. From Weeks 2 to 5, classroom instruction focused on developing coherence and cohesion in academic writing. Each week, students

submitted essays to ChatGPT and received feedback on aspects such as organization, linking devices, and clarity. They then revised and resubmitted their drafts based on the AI-generated feedback. In Week 6, participants completed a post-test essay and a perception questionnaire in relation to the usefulness, reliability, and relevance of ChatGPT feedback.

The quantitative data consisted of learners' pre- and post-test scores in the area of coherence and cohesion. The normality of the difference scores was first assessed using the Shapiro-Wilk test. As the data violated the assumption of normality, the scores were then analysed using the Wilcoxon signed-rank test to determine whether there was a statistically significant improvement in writing performance. Statistical significance was set at $p < .05$. The qualitative data derived from the Likert-scale questionnaire, which explored students' frequency of ChatGPT use and their perceptions of its feedback, including its reliability and usefulness.

4. Results

This section interprets the findings from the quantitative and qualitative analyses of the study. The primary aim was to examine how generative AI feedback improved IELTS learners' coherence and cohesion in writing. Data were drawn from three key sources: (a) pre- and post-writing task scores assessed based on the IELTS Writing rubric, (b) detailed language analyses of textual features associated with coherence and cohesion, and (c) learner perceptions collected through a questionnaire. The three complementary sources provided a broader perspective on the impact of generative AI, enabling a profound understanding of its role in fostering learners' writing development.

Table 1

Shapiro-Wilk test for difference scores

	W	p
CC Post-test - CC Pre-test	0.8259	.002

The normality of the difference scores (post-test minus pre-test) for learners' CC scores was examined using the Shapiro-Wilk test.

As can be seen in Table 1, the Shapiro-Wilk test indicated that the distribution of the difference scores significantly deviated from normality, $W = .8259$, $p = .002$. Because the p-value was less than .05 and IELTS Writing band scores are reported on an ordinal 9-band scale, the assumptions required for a paired-samples t-test were not fully satisfied.

According to Bluman (1997), the Wilcoxon signed-rank test was selected to evaluate the difference between the pre-test and post-test CC scores. This is because it does not require the assumption of normality and is appropriate for paired ordinal data.

Table 2

Descriptive statistics and Wilcoxon signed-rank test results for learners' pre-test and post-test writing performance

Measure 1		Measure 2	W	z	p	Rank-Biserial Correlation
CC Post-test	-	CC Pre-test	210.0	3.92	<0.001	1.00

The Wilcoxon signed-rank test showed a statistically significant improvement over participants' Coherence and Cohesion scores after the intervention, $W = 210.0$, $z = 3.92$, $p < .001$. Moreover, the matched rank-biserial correlation indicated an extremely large effect size ($rrb = 1.00$), suggesting that the ChatGPT-based feedback had a high positive impact on all of the learners' coherence and cohesion in writing.

After analysing the before and after versions of student essays, there was a great reduction in errors related to the use of cohesive devices. In the first drafts, cohesive devices were used both inconsistently and often inappropriately. For example, students often overused basic connectors such as “because,” “besides,” and “and,” which led to repetitive sentence structures and a lack of logical link. On the other hand, misuse or omission of linking phrases resulted in abrupt transitions between ideas, thus decreasing overall coherence. A common instance was the overuse of additive markers (e.g., “besides”) where contrastive or causal links were needed.

In the revised versions, however, students used cohesive devices more accurately and diversely. Such markers as “for instance,” “as a result,” “on the other hand,” and “therefore” were properly utilised to signal cause and effect, exemplification, as well as contrast. Hence, it allows for a clear and logical progression of ideas and well-developed paragraphs. For example, in the food waste essay, the transition “as a result” replaced the less precise “because,” thereby enhancing logical flow. Similarly, in the zoo essay, the use of “however” and “nevertheless” made sharper contrasts between opposing viewpoints.

In terms of quantitative research, the number of cohesive device errors, including inappropriate use, omission, or redundancy, significantly decreased in the revised essays. In the first drafts, errors in cohesion were frequent and sometimes hindered meaning comprehension, contributing to lower band scores, approximately Band 5.5 to 6.0 for coherence and cohesion. In the final drafts, cohesive devices were both more diverse and contextually appropriate, resulting in better ratings of around Band 7.0 to 7.5 in this category.

These findings indicate that explicit feedback and revision contributed to improved control over cohesion. Students not only could reduce errors in basic connectors but also demonstrated an ability to use a wider range of cohesive devices to express complex logical relationships. This development suggests greater awareness of how cohesive devices lead to coherence, a key criterion in the IELTS Writing Task 2 rubric.

The comparative analysis of the students’ before and after versions showed vast improvements in not only logical flow but also paragraphing. In the first drafts, paragraphs often contained loosely connected ideas associated with limited development. For example, in the food waste essay, several reasons for food wastage were listed continuously without clear sequencing or elaboration, causing repetition rather than progression. Similarly, the topic sentences in the essays on education and technology were vague, and supporting details were not linked back to the main argument; thus, the cohesion between sentences in a paragraph was weak.

In the revised drafts, the paragraph structure was clearer, consisting of stronger topic sentences, supporting arguments, and concluding statements. Each paragraph developed a single controlling idea, and transitions between paragraphs reflected a more logical sequencing of arguments. For instance, in the public transportation essay, the revised version began with the short-term benefits of road expansion before contrasting them with the long-term sustainability of public transport. This arrangement produced a more balanced discussion and complied with the requirements for coherence and cohesion of IELTS Writing Task 2.

Furthermore, logical flow within paragraphs improved through the use of appropriate discourse markers and clearer cause and effect reasoning. In the after versions, students moved beyond merely listing ideas and instead created a logical chain of reasoning. Some students not only provided the reasons to open a zoo but linked the ethical considerations of keeping animals in captivity with conservation function in a proper flow and transition words such as “however” and “on the other hand” were used effectively.

As a consequence, the students’ control over coherence and cohesion was improved as ideas were presented in a logical order and arguments were more fully developed. This improvement suggests that students became more aware of effective paragraphing strategies, such as starting with clear topic sentences, maintaining progression, using transitions to guide the reader, and drawing a reasonable conclusion. Based on IELTS Writing Task 2 descriptors, this shift is more likely to raise scores in the Coherence and Cohesion from approximately Band 5.5 to 6.0 in the first drafts to Band 7.0 to 7.5 in the revised versions. Automated textual analysis using Coh-Metrix (Graesser et al., 2014) confirmed these findings, showing improved measures of referential cohesion and deep cohesion in the learners’ writing.

Questionnaire is also conducted to gain insights into how students perceive of using ChatGPT to enhance coherence and cohesion in their writing.

Table 3

Student’s frequency of using AI to study English

Frequency of Use	**Percentage**
Every day	40%
Several times a week	50%
Occasionally	10%
Do not use	0%

When students were asked about how frequently they use AI tools to support their English learning, half of the students (50%) reported using AI several times a week, while about 40% used it daily. A smaller proportion (10%) indicated only occasional use, and none reported not using AI at all. These findings highlight the growing prevalence of AI technology among university students and its potential for integration into English teaching and learning.

Table 4

Students’ Self-evaluation of cohesion and coherence in their essay

Evaluation Level	**Percentage**
Good	5%
Fair	50%
Average	45%
Poor	0%

Students have also been asked to self-evaluate coherence and cohesion aspects in their academic essays. Most of them highly rated their ability as either average or fair, with very few considering themselves good, and almost none identifying as poor. Although students

expressed confidence in their skills, analysis revealed a wide range of errors related to cohesion and coherence in their first version of essay. Even though they manage to use some cohesive devices, they show their inability to use a variety of devices. This wrong self-evaluation suggests that students are oblivious of their writing deficiencies, which AI tools like ChatGPT can help illuminate.

Table 5

Sources that students usually receive feedback when learning to write in English

Source of Feedback	Percentage
No feedback received	0%
Peer review (from classmates)	50%
Guidebooks or reference materials	30%
AI tools (e.g., ChatGPT, Grammarly, etc.)	95%
Teacher	55%

Regarding the ways from which students get feedback to refine their writing essays, most of the students agree that AI tools are the most commonly used source (95%), followed closely by teachers and peer reviews (50%). No respondents reported receiving no feedback at all. The crucial role of technology, particularly AI tools, in enhancing writing skills has become increasingly evident. Students are gradually turning to AI not only for research and entertainment but also for educational purposes. This shift reflects a reduced dependence on teachers and a growing reliance on alternative sources of support in learning English.

Table 6

Time Students Spent Using Chat GPT to Edit Essays

Time	Percentage
Under 10 minutes	10%
10-20 minutes	60%
20-30 minutes	25%
More than 30 minutes	5%

In the experiment, when students were instructed to use Chat GPT to revise cohesion and coherence in their essays, more than half of them (60%) dedicated only about 10-20 minutes interacting with Chat GPT to finalize their ideas, exactly 25% of students spent between 20-30 minutes using ChatGPT. While 10% of students only seek advice from Chat GPT in less than 10 minutes, 5% of students spend over 30 minutes communicating with this AI tool.

Table 7

Purposes of using Chat GPT to enhance CC

Purposes	Percentage

Suggest ways to use references	50%
Suggest ways to use ellipses	35%
Suggest ways to use collocations	55%
Find appropriate collocations	55%
Find appropriate reiteration to avoid repetition	60%
Ask ChatGPT to detect and revise inappropriate linking errors	90%
Help check the logical flow of the article	60%
Suggest how to present content and illustrative images	55%
Improve accuracy and writing style	40%
Ask ChatGPT to compare different versions to choose the more logical one	60%

As for what purposes they use ChatGPT for in the process of editing their essays to achieve coherence and cohesion, the most common purpose for which ChatGPT is utilized is to "Ask ChatGPT to detect and revise inappropriate linking errors", with 90% respondents indicating this use. Finding appropriate collocations and reiteration to avoid repetition (60%) are the second top choices. "Ask ChatGPT to compare different versions to choose the more logical one" also gains acclaim from 60% of learners as they might be confused if they have different ways to develop ideas. Conversely, the least common purpose is suggesting ways to use ellipses, reported by only 35% of respondents. This suggests that students may undervalue or lack awareness of ellipsis as a cohesive device.

Table 8

Students' Attitudes Towards Utilizing Chat GPT to Enhance CC in Ielts Writing

		n	M	S.D.
1	I intend to continue using ChatGPT afterwards	20	4.10	0.72
2	I feel that Chat GPT's feedback is easy to understand & apply	20	4.00	0.79
3	I'm excited when using ChatGPT to improve writing	20	4.05	0.60
4	I believe ChatGPT makes me more confident	20	4.05	0.60
5	I'm encouraged by ChatGPT feedback	20	3.90	0.85
6	I am demotivated when feedback is unclear	20	3.10	1.17
7	I think using ChatGPT is boring and monotonous	20	3.05	1.05

The table illustrates student's attitude when seeking for Chat GPT's assistance in giving feedback on their cohesion and coherence in Ielts Writing. Positive feedback such as "I feel that Chat GPT's feedback is easy to understand & apply", "I intend to continue using

ChatGPT" and " I'm excited when using ChatGPT to improve writing", are approved by most of the correspondents, with their means are 4.00, 4.10 and 4.05 respectively. Notably, negative statements such as "ChatGPT is boring and monotonous" or "I feel demotivated when feedback is unclear" received relatively low levels of agreement.

Table 9

Student's perceptions of Relevance and Usefulness

	n	M	S.D.
Using ChatGPT saves me time when editing my writing	20	4.20	0.77
ChatGPT gives useful suggestions to improve linking between sentences/paragraphs using cohesive devices	20	4.20	0.62
ChatGPT helped me see whether my ideas are clearly organized and logically connected (coherence)	20	4.00	0.56
I am complete in control when revising based on feedback, using prompts and exploring coherence strategies	20	3.90	0.64
After using ChatGPT, I can apply cohesion/coherence strategies to other essays without help	20	3.85	0.75
ChatGPT's feedback is comparable to or better than peers/teachers in some cases	20	3.75	0.72
ChatGPT's feedback fits common contexts in IELTS Writing Task 2	20	3.70	0.80

Table 9 presents students' perceptions of the relevance and usefulness of ChatGPT in enhancing coherence and cohesion in writing. The data indicate that users highly value ChatGPT for its time-saving benefits (M = 4.20) and its helpful suggestions for improving cohesion through the use of linking devices (M = 4.20). Other aspects, such as its ability to help students examine the coherence of their writing (M = 4.00) and the opportunity to actively engage with ChatGPT using prompts to revise essays (M = 3.90), also received strong approval. However, perceptions regarding ChatGPT's independence are more reserved. Statements like "ChatGPT's feedback is comparable to or better than peers/teachers in some cases" (M = 3.75) and "After using ChatGPT, I can apply cohesion and coherence strategies to other essays without help" (M = 3.85) received comparatively lower ratings. These findings suggest that students view ChatGPT primarily as a supplementary tool rather than a full replacement for traditional instruction or peer feedback.

5. Discussion

The findings from the essay versions and questionnaire offer valuable insights into the efficacy of using ChatGPT as a tool for enhancing coherence and cohesion in academic writing. Notably, learners demonstrated a statistically significant improvement in coherence and cohesion following the intervention ($W = 210.0$, $z = 3.92$, $p < .001$), with a very large effect size (rank-biserial correlation = 1.00). This empirical progress directly supports Javahery's (2025) assertion that mastering logical progression and cohesive devices is essential to break the Band 6 threshold. These improvements also parallel previous research demonstrating that focused instruction and practice can significantly enhance learners' discourse competence and coherence (Weigle, 2013).

Post-test essays revealed clearer paragraphing, smoother transitions between ideas, and more varied and accurate use of cohesive devices compared to pre-test essays, reflecting real

progress toward higher IELTS band descriptor expectations (Taylor & Wigglesworth, 2009). The reduction in cohesion and coherence errors appears to stem from explicit, example-driven feedback provided by AI tools. Students' shift from overusing basic connectors to employing a broader range of cohesive devices indicates a deeper understanding of how these elements function in academic writing. This outcome shares similarities with Crossley et al. (2016), who argued that effective feedback on cohesion requires not merely pointing out errors but modeling alternative cohesive strategies, something generative AI was particularly skilled at delivering.

The questionnaire results further underscore the growing prevalence of AI tools in language learning, with most of the students reporting frequent use of such technologies to study English. This reflects a broader shift toward technology-driven learning methodologies. Students primarily used ChatGPT to detect and revise inappropriate linking errors and to find suitable collocations. In addition, some writing aspects, including linking ideas, organizing paragraph structure and better logical flow, have been reported to improve significantly by students. These improvements align with the findings of Sekewael and Anaktototy (2024) which advocate that feedback from ChatGPT can bolster coherence and cohesion in writing. One prominent merit of ChatGPT pointed out by the students is rapid response time, which is perceived as less time-consuming compared to traditional feedback methods from teachers or peers. Regarding future use, the majority of students expressed a positive intention to continue using ChatGPT as a scaffolding tool. However, some concerns were still raised about the accuracy of the feedback provided, with 14.3% of respondents expressing skepticism. This posits that the limitation of AI tools must be underscored and that students must apply critical thinking when engaging with AI feedback to ensure its relevance and reliability.

Even though the relevance and usefulness of ChatGPT affirms its role as a supplementary tool, the lower ratings concerning the independence of ChatGPT's feedback indicate that students still see traditional instruction and peer feedback as essential components of their learning journey. This suggests that while AI can enhance writing skills, it should complement rather than replace existing pedagogical practices.

6. Conclusion

This study sets out to examine the potential of generative AI feedback in enhancing coherence and cohesion in IELTS writing, a domain where candidates often struggle due to limited awareness of textual organization and logical flow. By investigating learners' engagement with AI-mediated feedback, the research provides valuable insights into how emerging technologies can augment traditional pedagogical practices in high-stakes second language (L2) writing contexts.

One of the central findings is that generative AI feedback, when applied thoughtfully, can support learners in addressing weaknesses in textual cohesion and overall logical structure. Unlike traditional forms of corrective feedback, which often focus on grammar, vocabulary, and surface-level errors (Ferris, 2010; Bitchener & Ferris, 2012), generative AI has the capacity to draw learners' attention to discourse-level issues, including transitions, thematic progression, and paragraph unity. This aligns with Hyland's (2005) assertion that successful academic writing requires more than accurate grammar; it requires the ability to construct coherent arguments and manage reader expectations effectively.

The findings also indicate that generative AI feedback fosters learner autonomy by providing immediate, individualised, and non-judgmental suggestions. Previous studies on automated writing evaluation (AWE) systems such as Grammarly (Dizon & Gayed, 2024) stressed the benefits of timely feedback, but these systems were often limited to fixed error

categories. In contrast, generative AI tools can simulate dialogic interaction and contextualise feedback within learners' specific writing tasks (Asadi et al., 2025). This adaptability allows learners to revise more strategically, thereby engaging in deeper processing of writing practices.

However, the integration of generative AI feedback is not without limitations. While learners benefited from guidance on coherence and cohesion, challenges arose in terms of overreliance and the potential for learners to adopt AI-generated suggestions uncritically. This concern resonates with broader debates in applied linguistics regarding the balance between scaffolding and learner agency (Ushioda, 2020; Warschauer & Liaw, 2011). Educators must therefore frame AI tools as complements, not substitutes, for critical engagement with writing. This highlights the importance of teacher mediation in helping students interpret and evaluate AI-generated feedback.

Future research should expand on these findings by examining the long-term effects of AI-assisted practice on IELTS Writing Task 2 performance, particularly whether improvements in coherence and cohesion transfer to other academic or professional writing contexts. In addition, it is important to investigate ethical considerations, such as the potential risk of standardising or synthesising learners' writing styles. Combining generative AI feedback with explicit writing instruction and peer collaboration may offer a balanced and effective approach to enhancing learners' writing development.

References

Anindita, W. K. (2024). Cohesion and Coherence Problems Among Non-Native English Students' Writing Essays. *Lensa: Kajian Kebahasaan, Kesusastraan, dan Budaya*, *14*(1), 58. https://doi.org/10.26714/lensa.14.1.2024.58-79

Asadi, M., Ebadi, S., & Mohammadi, L. (2025). The impact of integrating ChatGPT with teachers' feedback on EFL writing skills. *Thinking Skills and Creativity*, *56*, 101766. https://doi.org/10.1016/j.tsc.2025.101766

Bahang, M. D., Nawun, A. R., Wutun, A. A., & Nurhusain, M. (2021). The use of thematic progression in writing: Hortatory exposition text at second grade students of SMA Negeri 1 Elar. *The Educational Review*, *5*(10), 385–390. https://doi.org/10.26855/er.2021.10.002

Banerjee, J., & Wall, D. (2006). Assessing and reporting performances on pre-sessional EAP courses: Developing a final assessment checklist and investigating its validity. *Journal of English for Academic Purposes,* *5*(1), 50-69. https://doi.org/10.1016/j.jeap.2005.11.003

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback?. *International Journal of Educational Technology in Higher Education*, *21*(1), 23. https://doi.org/10.1186/s41239-024-00455-4

Behbahani, H. K., Jabbari, A. A., & Farahani, A. H. D. (2018). Effectiveness of explicit instruction of cohesive devices on Iranian EFL learners' writing development. *International Journal of Advanced Multidisciplinary Scientific Research*, *1*(4), 1-10. https://doi.org/10.31426/ijamsr.2018.1.4.211

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge. https://doi.org/10.4324/9780203832400

Bluman, A. G. (1997). *Elementary statistics: A step by step approach* (3rd ed.). McGraw-Hill.

Chen, M., & Cui, Y. (2022). The effects of AWE and peer feedback on cohesion and coherence in continuation writing. *Journal of Second Language Writing*, *57*, 100915. https://doi.org/10.1016/j.jslw.2022.100915

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior research methods*, *48*(4), 1227-1237. https://doi.org/10.3758/s13428-015-0651-7

Diep, G. L., & Le, T. N. D. (2024). An Analysis of Coherence and Cohesion in English Majors' Academic Essays. *International Journal of Language Instruction*, *3*(3), 1-21. https://doi.org/10.54855/ijli.24331

Dizon, G., & Gayed, J. M. (2024). A systematic review of Grammarly in L2 English writing contexts. *Cogent Education*, *11*(1), 2397882. https://doi.org/10.1080/2331186X.2024.2397882

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition, 32*(2), 181-201. https://doi.org/10.1017/S0272263109990490

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, *115*(2), 210-229. https://doi.org/10.1086/678293

Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in English*. Routledge.

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated argumentative student essays. *Scientific Reports, 13*(1), 18617. https://doi.org/10.1038/s41598-023-45644-9

Hyland, K. (2003). *Second language writing*. Cambridge University Press.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83-101. https://doi.org/10.1017/s0261444806003399

Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (2nd ed., pp. 1-22). Cambridge University Press. https://doi.org/10.1017/9781108635547.003

Janebi Enayat, M. (2025). Computationally derived linguistic features of L2 narrative essays and their relations to human-judged writing quality. *Language Testing in Asia, 15*(1), 35. https://doi.org/10.1186/s40468-025-00374-9

Javahery, P. (2025). *Analyzing Discourse Patterns in ChatGPT-Generated IELTS Writing Task 2 Essays: A Discourse Analysis Approach*. https://doi.org/10.21203/rs.3.rs-5608928/v1

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Kormos, J. (2023). The role of cognitive factors in second language writing and writing to learn a second language. *Studies in Second Language Acquisition*, *45*(3), 622-646. https://doi.org/10.1017/S0272263122000481

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, *54*(2), 537-550. https://doi.org/10.1177/00336882231162868

Leli, L. (2020). Analysis of Coherence and Cohesion on Students' Academic Writing: A Case Study at The 3rd Year students at English Education Program. *Alsuna: Journal of Arabic and English Language, 3*(2), 74-82. https://doi.org/10.31538/alsuna.v3i2.980

Li, X., & Chen, M. (2026). Assessing the effects of explicit coherence instruction on EFL students' integrated writing performance. *Assessing Writing*, *67*, 101019. https://doi.org/10.1016/j.asw.2026.101019

Li, S., Prior, M., Nero, S., Hiver, P., Al-Hoorie, A. H., Murakami, A., Wei, L. , & Ortega, L. (2023). Methodological innovation in applied linguistics research: Perspectives, strategies, and trends. *Language Teaching*, *56*(4), 551-556. https://doi.org/10.1017/S026144482300023X

Oshima., Alice., & Hogue, A. (2006). *Writing academic English* (4th ed.). Pearson Education.

Pavavijarn, S. (2022). Influences of thematic progression on quality of EFL argumentative writing. *LEARN Journal: Language Education and Acquisition Research Network, 15*(1), 282-319.

Rahman, G., Zaigham, M. S., & Umer, M. (2023). A study of cohesive devices in students' academic writing. *Qlantic Journal of Social Sciences and Humanities*, *4*(4), 372-384. https://doi.org/10.55737/qjssh.897009451

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning, 53*(1), 165-202. https://doi.org/10.1111/1467-9922.00213

Sekewael, M., & Anaktototy, K. (2024). Chat-GPT Feedback: Fostering Growth in English Writing Abilities. *Interdisciplinary Themes of Sociolinguistic Studies-Language Learning, Technology, and Identity in the Age of AI and Globalization: Language Learning, Technology, and Identity in the Age of AI and Globalization.* IntechOpen. https://doi.org/10.5772/intechopen.1007287

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, *36*(2), 187-209. https://doi.org/10.1017/S0958344023000265

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, *91*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Sun, Q., Chen, F., & Yin, S. (2023). The role and features of peer assessment feedback in college English writing. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1070618

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075. https://doi.org/10.1016/j.caeai.2022.100075

Taylor, L., & Wigglesworth, G. (2009). Are two heads better than one? Pair work in L2 assessment contexts. *Language Testing, 26*(3), 325-339. doi.org/10.1177/0265532209104665

Tengler, K., & Brandhofer, G. (2025). Exploring the difference and quality of AI-generated versus human-written texts. *Discover Education*, *4*(1), Article 209. https://doi.org/10.1007/s44217-025-00529-z

Termjai, M. (2024). Peer Feedback in Thai EFL Writing: Students' Perceptions, Accuracy, and Revisions. *English Language Teaching*, *17*(8), 10. https://doi.org/10.5539/elt.v17n8p10

Test taker performance 2024-2025. (n.d.). IELTS. Retrieved August 29, 2025, from https://ielts.org/researchers/our-research/test-statistics

Trang, N. T. B., Nhu, N. V. Q., & Tuan, N. Q. (2023). Use of cohesive devices in paragraph writing by EFL students at English language centers in Vietnam. *VNU Journal of Foreign Studies, 39*(3), 152-169. https://doi.org/10.63023/2525-2445/jfs.ulis.5099

Ushioda, E. (2020). *Language learning motivation: An ethical agenda for research.* Oxford University Press.

Warschauer, M., & Liaw, M.-L. (2011). Emerging technologies for autonomous language learning. *Studies in Self-Access Learning Journal, 2*(3), 107-118.

Weigle, S. C. (2013). English language learners and automated scoring of essays: Critical considerations. *Assessing Writing, 18*(1), 85-99. https://doi.org/10.1016/j.asw.2012.10.006

Bionote

Hoàng Ngọc Quỳnh Như is an English language lecturer at the Industrial University of Ho Chi Minh City, Vietnam. She holds an MEd in TESOL from Edith Cowan University. She specialises in developing students' speaking, writing, listening, and reading skills. Her research focuses on applying Al-based tools to enhance English proficiency among university students, exploring attitudes and practical solutions. She has also conducted studies on pedagogical strategies, language assessment, collaborative learning, and the use of technology in teacher education.

Co- Trần Sĩ Nguyên Sa is a visiting lecturer in English at HCMC University of Technology, Vietnam. She holds an MEd in TESOL Presenters' from Edith Cowan University. She focuses on improving students' proficiency in speaking, writing, listening, and reading. Her research interests include Al applications in language education, digital literacy, and the sociocultural impact of technology on language teaching and learning. She also explores innovative pedagogical strategies and the role of technology in teacher training and curriculum development.

#27 Evaluating the Impact of Vocabulary Learning through Online Websites and Digital Resources on University EFL Students: Teachers' Perspectives

Nguyễn Thuỷ Hương[1]

Abstract

The rapid growth of digital technology has transformed teaching vocabulary in English as a Foreign Language (EFL), university students have begun employing online resources to study words on their own. In Vietnam, vocabulary websites, online dictionaries and interactive platforms provide learners a helpful substitute for classroom instruction for many learners. Teachers' perspectives on the educational usefulness of these tools are, however, not explored extensively. This investigation attempts to fill that gap by studying Vietnamese university EFL teachers' perceptions of the efficacy, advantages and limitations of internet vocabulary learning materials. With a convergent parallel mixed methods approach, quantitative data was obtained using questionnaires from 60 teachers; data were derived from 12 semi-structured interviews with academics. Descriptive and correlational analyses revealed macro-behavioural trends in general and thematic analysis investigated teachers' perceptions. Results showed that teachers regarded online resources to be supportive for increasing student vocabulary, independence and motivation. However, issues include inconsistent learning behavior, poor digital literacy, and poor self-regulation among students. The authors found that teachers have a pedagogical influence when implementing digital tools and monitoring students' progress and connection between online activities and curricular aims in a successful web learning scenario, in the end the effective use of technology in a web learning activity to students will not be possible without careful instruction and not with the use of technology alone.

Keywords: web-based vocabulary learning; EFL; teachers' perspectives; mixed-methods research; Vietnamese higher education

Tóm tắt

Sự phát triển nhanh chóng của công nghệ số đã làm thay đổi cách dạy và học từ vựng trong lĩnh vực học tiếng Anh như một ngoại ngữ (EFL), khi sinh viên đại học ngày càng sử dụng các nguồn tài nguyên trực tuyến để tự học từ vựng. Tại Việt Nam, các trang web học từ vựng, từ điển trực tuyến và các nền tảng tương tác trở thành những công cụ hỗ trợ hữu ích, giúp bổ sung cho việc học trên lớp. Tuy nhiên, quan điểm của giáo viên về giá trị sư phạm của những công cụ này vẫn chưa được nghiên cứu đầy đủ.

Nghiên cứu này nhằm lấp đầy khoảng trống đó bằng cách tìm hiểu nhận thức của giáo viên đại học tại Việt Nam dạy tiếng Anh như một ngoại ngữ về hiệu quả, lợi ích và hạn chế của các tài liệu học từ vựng trên Internet. Sử dụng phương pháp hỗn hợp song song hội tụ, dữ liệu định lượng được thu thập thông qua phỏng vấn với 60 giáo viên, trong khi dữ liệu định tính được thu thập từ 12 cuộc phỏng vấn bán cấu trúc. Phân tích mô tả, tương quan và phân tích chủ đề được áp dụng để khám phá xu hướng và quan điểm của giáo viên.

Kết quả cho thấy giáo viên đánh giá các nguồn trực tuyến giúp cải thiện vốn từ vựng, tính tự chủ và động lực học tập của sinh viên. Tuy nhiên, vẫn tồn tại các vấn đề như thói quen học tập thiếu nhất quán, kỹ năng số hạn chế và khả năng tự quản lý thấp. Nghiên cứu nhấn mạnh vai

[1] *University of Greenwich; Email: huongnt114@fe.edu.vn*

trò hướng dẫn sư phạm của giáo viên trong việc tích hợp công nghệ, theo dõi tiến độ và liên kết hoạt động trực tuyến với mục tiêu chương trình học, bởi việc sử dụng công nghệ chỉ thực sự hiệu quả khi có sự định hướng giảng dạy cẩn trọng.

Từ khóa: học từ vựng dựa trên web; EFL; quan điểm của giáo viên; nghiên cứu phương pháp hỗn hợp; giáo dục đại học Việt Nam.

1. Introduction

Knowledge of vocabulary is a vital component in second and foreign language development, which serves as the foundation for learners' understanding of input, accurate expression of ideas and the successful participation in academic and communicative tasks. In particular to university EFL contexts, students must deal with progressively more complex texts and discourse focused on the disciplines in which they find themselves, which places considerable demands on their lexical competence. As a result, the teaching field of vocabulary has emerged as one of the key challenges facing English language education, especially when classroom time is restricted and learner proficiency is varied. Vocabulary learning has been transformed by the recent advancements of digital technology, which are opening the door to online resources that extend learning beyond classroom walls. Multimodal input, interactive exercises, and opportunities for repeated exposure and retrieval are included in online platforms, vocabulary learning websites, and mobile-assisted apps that support learning. Research has empirically indicated that technology-enabled vocabulary learning contexts can lead to significant gains in vocabulary acquisition and retention compared to traditional practices (Zou et al., 2021; Şimşek & Şimşek, 2025). These benefits are generally linked to higher student engagement, the change from one pace of learning with a set vocabulary at different steps, and deeper lexical processing. However, pedagogical affordances have not ensured success of web-based vocabulary learning. However, there is increasing attention in studies that technology by itself does not cause learning to have meaningful consequences; it is mediated through pedagogic design, learner strategies, and instructional support (Ertmer & Ottenbreit-Leftwich, 2010). Teachers in particular are critical in determining which of the web resources are chosen, used and related to teaching, learning and evaluation. Teachers' attitudes towards technology, their understanding of digital technology competence, and the students' readiness to learn about technology have a bearing on the level and significance of technology integration in a language classroom. The use of web-based vocabulary resources has become more prevalent in EFL classrooms in Asia and in Vietnam in particular in recent years, and with the rising prominence of online, blended learning technologies. As learners turn more and more to online websites and Internet resources for independent vocabulary learning, however, not much is known about how teachers view the success and struggle of using these resources empirically. Most previous research has concentrated on learner outcomes, failing to explore teachers' perspectives, especially in higher education and context-specific settings like Vietnam ((Rahmani et al., 2022; Guo et al., 2022). To fill this gap, the current research seeks to investigate Vietnamese university EFL teachers' views on the impact of web-based vocabulary learning resources on students' vocabulary development. By investigating teachers' perceived benefits, difficulties and pedagogical issues, this investigation seeks to provide insights into ways in which web-based learning of vocabulary may be well incorporated in EFL teaching at tertiary level. In addition to filling an empirical void, this research contributes to research and practice on EFL vocabulary teaching from many angles. First is that it moves the emphasis of analysis from outcome-orientated to teachers as central mediators of web-based vocabulary learning, to the extent that it looks at how pedagogical beliefs and pedagogic choice influence the efficacy of digital resources. Second, the fact that this study uses a mixed-

methods design means that the study is both broad and deep, and it provides qualitative pedagogical perspectives to put quantitative trend in perspective. Third, the research adds context specific evidence of Vietnam, a rapidly developing EFL context with no published evidence internationally exploring digital vocabulary learning. Together, these contributions advance the current understanding of technology-enhanced vocabulary pedagogy and contribute to more contextually sensitive instructional practices.

2. Literature

2. 1. Theoretical Foundations of Web-Based Vocabulary Learning

Web-based vocabulary learning is theoretically based on several influential frameworks in applied linguistics and educational psychology. Theoretically, learning vocabulary is more effective when people are involved in processing the lexical items through multiple cognitive operations, including semantic elaboration, contextualization, and repeated retrieval (depth of processing theory). Such systems are often supported in digital vocabulary resources via multimodal input, interactive tasks, and spaced repetition mechanisms. Moreover, sociocultural theory highlights the mediating role of tools in learning, framing the usage of web-based platforms as mediational means that shape learners' interaction with their lexical knowledge. Online vocabulary environments enable learners to negotiate meaning, have access to authentic examples of vocabulary use, and receive immediate feedback, which is beneficial for the internalization of vocabulary knowledge. These theoretical perspectives collectively support the pedagogical potential of web-based vocabulary learning when it is meaningfully integrated into instructional contexts.

2.2. Vocabulary Learning Outcomes in Technology-Mediated Environments

An increasing number of empirical studies have examined vocabulary learning outcomes in technology-mediated settings. The meta-analysis suggests that technology-assisted vocabulary instruction produces significantly higher vocabulary gains than traditional strategies, with its impact in higher education contexts being most significant (Boulton & Cobb, 2017; Zou et al., 2021). This is due to greater exposure, learner control, and numerous opportunities to practice the material multiple times. Evidence from the field of digital reading environments also suggests that glossed texts, hyperlinked vocabulary, and interactive annotations are positively related to incidental and intentional acquisition of vocabulary. Those exposed to digitally enhanced texts have better vocabulary retention and form–meaning mapping compared to those learning with non-digital ones (Guo et al., 2022). More importantly, studies also show that the gains in vocabulary learning are even greater when students are explicitly instructed about digital resources and how to use them strategically.

2.3. Learner Autonomy and Self-Regulated Vocabulary Learning

Web-based vocabulary learning encourages learner autonomy and self-regulated learning to a large degree. Digital platforms help to set learning goals, manage learner progress and adjust learning strategies to suit varying individual learner needs. For instance, there are empirical findings showing that learners with greater self-regulated behavioural engagement using digital vocabulary tools experience more vocabulary growth and retention (Dornyei & Ryan, 2015; Guo et al., 2022). But technology access does not automatically lead to autonomy. However, without pedagogical scaffolding, researchers see that learners may resort to surface-level strategies (e.g., rote memorization of word lists), reducing the opportunities for long-term

vocabulary development. This finding confirms the importance of teacher guidance in cultivating effective self-regulated vocabulary learning behaviors in web-based environments.

2.4. Teachers' Roles in Mediating Web-Based Vocabulary Learning

Although learners engage with web-based vocabulary resources, teachers are key agents mediating the effectiveness of their pedagogy. Research on technology integration reveals that teachers' beliefs, technological competence, and instructional intentions underpin how digital tools are incorporated in classroom practice (Ertmer & Ottenbreit-Leftwich, 2010). In vocabulary instruction, teachers play an influencing role not only in tool selection but also in task design, learning objectives, and assessment alignment. Empirical research investigating teachers' perspectives shows that although educators recognize the motivational and instructional value of web-based vocabulary tools, teachers are concerned about monitoring learner engagement and maintaining depth of learning (Latip & Mahadi, 2021; Katemba, 2020). Of particular relevance to higher education settings are these issues, where independent learning is promoted and teacher monitoring is limited outside classroom settings.

2.5. Contextual Factors in Asian and Vietnamese EFL Settings

The effectiveness of web-based vocabulary learning has been significantly mediated by contextual factors in EFL contexts. Academic work performed in Asian settings underscores the contribution of institutional expectations, examination-oriented curricula, and technological infrastructure on teachers' willingness and competence in integrating digital tools (Hwang et al., 2008). Factors such as cultural beliefs about teachers' authority and learner independence also influence the use of web-based learning in pedagogical practices. Research on technology-enhanced language learning in Vietnam indicates that although teachers see growing pedagogical value in using digital resources, challenges related to training, workload, and curricular alignment still exist (Trinh et al., 2022). There is a notable dearth of studies systematically investigating teachers' views about web-based vocabulary learning using mixed-methods designs, indicating a clear gap for research in this context.

2.6. Synthesis and Research Gap

Taken together, the literature offers strong indicators regarding the impact of web-based vocabulary learning on lexical development, learner autonomy, and engagement in EFL settings. Several gaps remain, however. First, teachers' perspectives are relatively underreported compared to those of studies focused on learners. Second, there is a paucity of literature in the field of web-based vocabulary learning in Vietnamese institutions of higher education using robust mixed-methods designs. Lastly, how teachers' beliefs and pedagogical practices interact with perceived learning outcomes should be further examined. Filling these gaps is also important for advancing contextually responsive and pedagogically sound web-based vocabulary instruction.

Research Questions

Based on the reviewed literature, the study seeks to address the following research questions:

1. How do university EFL teachers in Vietnam perceive the effectiveness of web-based vocabulary learning resources for supporting students' vocabulary development?
2. What benefits and challenges do teachers identify in the implementation of web-based vocabulary learning resources in university EFL contexts?

3. How do teachers' perceptions influence their pedagogical practices and expectations regarding students' vocabulary learning outcomes?

3. Methodology

3.1. Research Design

Research questions were complemented with qualitative data; therefore, the study adopted a mixed-methods research design. A mixed-methods design was chosen because it allows for the integration of quantitative trends and qualitative insights in order to gain richer insights into teachers' perspectives concerning web-based vocabulary learning resources. More specifically, a convergent parallel design was used to collect quantitative and qualitative data simultaneously, analyze them independently, then integrate their data during interpretation. The quantitative part captured the overall patterns of teachers' beliefs for the effectiveness, benefits, and challenges of using web-based vocabulary learning resources (RQ1 and RQ2). The qualitative aspect aimed to obtain a broader insight to teachers' experiences and rationale for instruction such as their perceptions that shaped instruction and their expectations related to achieving student outcomes (RQ2 and RQ3).

3.2. Pedagogical Setting

The study was conducted within the Vietnamese higher education system, where English is taught as a foreign language and is used for both general and academic purposes. At the university level, English as a Foreign Language (EFL) education usually includes mandatory English courses for students who are not majoring in English, as well as specialized language courses for those who are majoring in English. Vocabulary learning is often integrated with reading, speaking, and writing classes, with less time dedicated solely to vocabulary development in class. In recent years, many universities in Vietnam have encouraged or subtly supported the use of web-based vocabulary learning tools, such as online dictionaries, vocabulary websites, spaced repetition applications, and mobile phone-enabled learning platforms. However, the extent to which these resources are integrated into teaching depends on individual teachers' pedagogical beliefs and technological skills.

3.3. Participants

Among EFL teachers at public and private universities in Vietnam, all participants had prior experience using or recommending web-based vocabulary learning resources in their teaching.

Table 1. Participant Demographic Information

Variable	Category	n	Percentage (%)
Gender	Female	38	63.3
	Male	22	36.7
Teaching Experience	1–5 years	15	25.0
	6–10 years	21	35.0
	Over 10 years	24	40.0
Academic Qualification	MA	44	73.3
	PhD	16	26.7
Teaching Context	English majors	27	45.0
	Non-English majors	33	55.0

In all, 60 teachers completed the questionnaire. Out of them, 12 teachers were purposively selected for follow-up semi-structured interviews in order to ensure diversity in teaching experience, institutional context, and frequency of technology use.

3.4. Instruments

3.4.1 Questionnaire

A structured questionnaire was developed to collect quantitative data on teachers' perceptions. The questionnaire consisted of three main sections aligned with the research questions:

Section A: Perceived effectiveness of web-based vocabulary learning resources (RQ1)

Section B: Perceived benefits and challenges of implementation (RQ2)

Section C: Pedagogical practices and expectations regarding student vocabulary outcomes (RQ3)

Items were measured using a five-point Likert scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). The questionnaire was reviewed by two experts in EFL pedagogy to ensure content validity and clarity. A pilot study (n = 15) was conducted, resulting in a Cronbach's alpha coefficient of 0.87, indicating high internal consistency.

3.4.2 Semi-Structured Interviews

Semi-structured interviews were carried out in order to gather qualitative information and insights into teachers' experiences and beliefs. The interview structure consisted of open-ended questions on the:

Teachers' evaluations of students' vocabulary development through web-based resources

Observed changes in learner autonomy and engagement

Challenges in guiding and monitoring students' use of online tools

Pedagogical decision-making related to technology integration

Interviews lasted approximately 30–45 minutes and were audio-recorded with participants' consent.

3.5. Data Collection Procedures

Data were collected over a 12-week period. The questionnaire was distributed online via institutional email and professional teaching networks. Participation was voluntary, and anonymity was assured. Interviews were conducted after the questionnaire phase to allow preliminary quantitative findings to inform qualitative exploration. All interviews were transcribed verbatim for analysis.

3. 6. Data Analysis

3.6.1 Quantitative Data Analysis

Statistical software was used to analyze quantitative data. Descriptive statistics, including means, standard deviations, and frequencies, were computed to summarize teachers' perceptions.

Table 2. Teachers' Perceptions of Web-Based Vocabulary Learning Resources

Item	Mean	SD
Enhances students' vocabulary knowledge	4.21	0.63
Promotes learner autonomy	4.35	0.58
Increases student motivation	4.18	0.66
Requires structured teacher guidance	4.47	0.54
Difficult to monitor students' usage	3.89	0.71

Inferential analyses (e.g., correlation analysis) were conducted to examine relationships between teachers' perceptions and reported pedagogical practices (RQ3).

3.6.2 Qualitative Data Analysis

Qualitative data were analyzed using thematic analysis following an iterative coding process. Initial codes were generated from the interview transcripts and subsequently grouped into broader themes. To enhance reliability, a second coder independently reviewed a subset of transcripts, with discrepancies resolved through discussion.

Table 3. Summary of Qualitative Themes

Theme	Description
Perceived pedagogical value	Teachers' beliefs about vocabulary gains and learning efficiency
Learner autonomy and responsibility	Students' self-directed use of online resources
Instructional challenges	Monitoring, guidance, and time constraints
Teacher mediation	Strategies for scaffolding and aligning tools with curriculum

2.7. Integration of Quantitative and Qualitative Findings

Both sets of data were integrated during interpretation to address the research questions holistically. Quantitative results provided an overview of general perception trends, while qualitative insights explained underlying pedagogical reasoning and contextual constraints.

4. Results

4. 1. Quantitative Results

The quantitative analysis suggested that the perceptions held by Vietnamese university EFL teachers towards web-based vocabulary learning resources were relatively positive. According to Table 2, teachers agreed with each other that these resources contribute to students' vocabulary knowledge development (M = 4.21, SD = 0.63) and are supportive towards autonomy in learners (M = 4.35, SD = 0.58). The perceived enhancement regarding student motivation was also relatively high (M = 4.18, SD = 0.66), suggesting that web-based

tools were perceived by teachers as helpful affective and cognitive agents to support vocabulary learning. Interestingly, the item with the highest mean score is web-based vocabulary learning requiring structured teacher mediation (M = 4.47, SD = 0.54). This point indicates a considerable agreement among teachers that digital learning platforms are inadequate in themselves without pedagogical scaffolding. Teachers reported that monitoring students' independent usage of web-based resources (M = 3.89, SD = 0.71) was moderate, as they felt the teaching process posed challenges (as opposed to outright resistance). Correlation analysis then showed a significant positive correlation between teachers' perceptions of effectiveness and the frequency with which they reported integrating web-based vocabulary tools into instruction ($r = .42$, $p < .01$), suggesting that favorable beliefs were associated with greater pedagogical use.

2.2. Qualitative Results

Thematic analysis of interview data led to the conclusion that there were four general themes, which are summarized and elaborated in Table 3. First, in terms of perceived pedagogical value, teachers perceived observed growth in vocabulary range, especially in receptive knowledge and lexical recognition. Some of the participants reported confidence when encountering new vocabulary in reading and listening tasks. Second, the central tenet of learner autonomy and responsibility reinforced teachers' observations regarding the extent to which web-based resources prompted students to be take-charge of vocabulary learning. However, autonomy was typically noted to be uneven, with more proficient or motivated students experiencing more consistent benefits than lower-level learners. Third, instructional challenges were a recurring theme. Teachers worry about students' reliance on superficial learning strategies like memorizing single word meanings without meaningful context. Difficulty tracking students' real use of online tools outside of class was also often cited. Finally, the teacher mediation theme pointed out the importance of guiding. Teachers described implementing strategies, such as suggesting websites, assigning learning activities, and incorporating online vocabulary activities into assessments, to ensure purposeful use of web-based resources.

5. Discussion

5. 1. Teachers' Perceptions of Effectiveness (RQ1)

Results suggest that Vietnamese university EFL teachers generally perceive the use of web resources for vocabulary learning as facilitating students' vocabulary growth. The high mean values of vocabulary enrichment and learner empowerment correlate with previous literature indicating that digital tools offer flexible, learner-centered learning platforms for vocabulary. Crucially for us, the findings imply that not only are teachers aware of cognitive advantages, they also recognize motivational ones, and this supports that web-based resources are useful adjuncts to EFL instruction.

5.2. Perceived Benefits and Challenges (RQ2)

Although teachers saw a number of benefits, the results also suggest that there was a nuanced understanding of the challenges. Notably, teachers appeared to have not completely viewed web-based vocabulary learning as a fully self-sufficient method, with considerable consensus in favor of providing structured guidance. Rather, they see it as best suited to a pedagogical framework that supports strategic learning and contextualized use of vocabulary. The moderate level of difficulty reported for monitoring the usage of students is indicative of a common issue in technology-enhanced learning contexts, especially in higher education contexts where independent learning is emphasized. These findings indicate that these

challenges are not intrinsic to the digital technologies but arise due to instructional design, learner training, and accountability related problems.

5.3. Pedagogical Practices and Teacher Mediation (RQ3)

The positive relationship between teachers' beliefs and their instructional practice reflects the impact of teacher beliefs on classroom application. Teachers who perceived web-based vocabulary resources as advantageous were more likely to systematically incorporate them into instruction. The qualitative results also highlight the role of teacher mediation by suggesting that task design, support, and alignment with assessments can be used to scaffold student use of digital tools. This finding reinforces the argument that teachers remain central agents in technology-enhanced vocabulary learning. Rather than replacing traditional instruction, web-based resources function most effectively as pedagogically mediated supplements that extend learning opportunities beyond the classroom.

5.4. Implications for EFL Vocabulary Instruction in Vietnam

Collectively, the findings indicate that the effectiveness of web-based vocabulary learning resources in Vietnamese higher education depends largely on the pedagogical approach taken. Positive beliefs and perceived problems of teachers also indicate the requirement of professional development on digital pedagogy and vocabulary learning strategies. Institutional support, in the form of training and curriculum alignment, might enable the sustainable use of web-based vocabulary resources in EFL instruction.

6. Conclusion

Through a mixed-methods approach, this study explored how web-based vocabulary learning resources affected university EFL students in Vietnam from the teachers' perspectives by examining both broad perceptual trends and subtle pedagogical information. The research integrates quantitative questionnaire data with the interviews of Vietnamese EFL teachers to provide a robust understanding of how Vietnamese EFL teachers perceive both the pedagogical value and drawbacks of web-based vocabulary learning activities at institutions of higher education. As a result, the results show teachers perceive the potential use of web-based vocabulary resources with broadly positive attitudes. Teachers found these digital tools effective for helping learners develop new vocabulary in the area of increasing lexical knowledge, target language input, and opportunities to practice outside of the classroom (repeated practice). Web-based resources were also considered as promoting learner autonomy and fostering student motivation because of the control learners have over learning pace, material and learning methods. These impressions indicate an ability for web-based vocabulary education resource as a complement to in class instruction in promoting vocabulary skills into flexible and learner-focused settings. At the same time, the results show that there are common understandings among teachers that the practice of web-based vocabulary learning has no inherent self-regulation or pedagogic sufficiency in isolation. In fact, teachers repeatedly highlighted the importance of having guided instruction to support vocabulary learning that is meaningful and strategic and not in the form of rote memorization or an individualized approach to learning vocabulary. There are evident problems – such as tracking students' independent use of online resources, differences in the digital literacy and self-regulation skills of learners, or fluctuations in learning strategies – and those highlight the shortcomings of unguided, unorganized use of digital vocabulary tools. These shortcomings highlight that the success of web-based vocabulary learning depends on the pedagogics of framing and supporting these resources. Crucially, this research highlights the importance of teachers as intermediaries in technology-mediated vocabulary learning. We conclude that the beliefs,

expectations, and teaching practice of teachers are significant to the choice, integration, and congruence of web-based vocabulary resources with curricular aims. Instead of being seen as a displace to traditional vocabulary teaching methods, web-based resources were found to be effective when part of a coherent pedagogical framework (comprising explicit instructions, task design, and mapping on to assessment procedure) as well. These results emphasize that technology offers an enabling device whose utility is the result of sound pedagogical use. By emphasizing the teachers' point of view in a Vietnamese university educational setting, this paper has a significant implication on the research on technology-supported vocabulary acquisition in EFL contexts. It fills an important gap in research that has been mainly concerned with the impact a student's vocabulary production has, which has not focused enough on teachers' instructional reasoning and contextual constraints. Based on quantitative and qualitative evidence, the study provides contextually relevant insights into the opportunities and challenges posed by internet language learning in Vietnam. It contributes significantly both to current literature on learning and EFL and to research on potential for the use of online vocabulary materials in sustainable and effective pedagogical practices to support vocabulary learning of university students.

7. Education and implications

The implications of this study are notable, with substantial significance for EFL vocabulary instruction, teacher professional development, and institutional policy in higher education. First, use web-based vocabulary learning resources pedagogically rather than as stand-alone tools. Teachers are instructed to explicitly support students in choosing suitable online resources, effective vocabulary-learning strategies, and applying newly acquired vocabulary in relevant settings. Structured tasks, including guided online vocabulary activities and following up with applying the learning in class, can serve as an antidote to superficial learning. Second, the findings emphasize the necessity of teacher education on digital vocabulary instruction. More pedagogical professional learning should include but are not limited to technical development such as instructional design, learner scaffolding, and evaluation for web-based vocabulary learning. Improving teachers' digital pedagogical competence might result in more assured and effective incorporation of online resources. The third must be structured to support learner autonomy. Web-based resources can encourage individual learning, though this often requires a clear explanation and explicit teaching of how to use those tools strategically to achieve meaningful learning impact. It is expected that through metacognitive training and reflection students will develop better and more sustainable vocabulary learning behaviour. Lastly, at an institutional level, both curriculum architects and faculty should address how they connect web-based vocabulary learning to course objectives and assessment strategies. In some ways, the institutional support which includes the proposed tools and clear guidelines for use and access to reliable digital resources will be necessary to support more widespread and fair implementation to an extended range of teaching contexts. Ultimately, web-based vocabulary learning resources are shown to offer significant potential for improving EFL vocabulary instruction in Vietnamese tertiary education. But the effectiveness of the latter depends greatly on well-informed pedagogical integration, teacher continuing mediation, and institutional support. It is hoped that future research could further this line of inquiry by assessing longitudinal learning outcomes, student perspectives, and comparative instructional approaches from various educational contexts.

References

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393.

Dornyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, *42*(3), 255–284.

Guo, F., Zhang, Y., & Wu, Z. (2022). Mobile-assisted vocabulary learning through the Shanbay App outside the classroom: Effects of self-regulation and peer scaffolding. *Frontiers in Psychology*, *13*, 993224.

Hwang, G.-J., Tsai, C.-C., & Yang, S. J. H. (2008). Criteria, strategies and research issues of context-aware ubiquitous learning. *Journal of Educational Technology & Society*, *11*(2), 81–91.

Katemba, C. V. (2020). Teachers' Perceptions in Implementing Technologies in Language Teaching and Learning. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, *5*(2), 123–136.

Latip, S. N. H. M., & Mahadi, I. (2021). *Teachers' Perceptions in Using Online Applications to Teach Vocabulary During Online Learning*.

Rahmani, A., Asadi, V., & Xodabande, I. (2022). Using mobile devices for vocabulary learning outside the classroom: Improving the English as foreign language learners' knowledge of high-frequency words. *Frontiers in Psychology*, *13*, 899885.

Şimşek, E. E., & Şimşek, B. (2025). Technology-assisted vocabulary learning for preschool children: a meta-analysis. *Computer Assisted Language Learning*, 1–33.

Trinh, T. H., Nguyen, M. N., & Tran, T. T. H. (2022). Teachers and Students' Perceptions of Using Digital Games in Improving Vocabulary at Non-English-majored Class. *AsiaCALL Online Journal*, *13*(5), 112–131.

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning*, *34*(5–6), 751–777.

ABOUT THE AUTHOR

Ms. Nguyen Thuy Huong (M.A.) is a dedicated educator and researcher with over 15 years of pedagogical experience within the FPT Education organization. She currently serves as a lecturer at FPT University, Greenwich Vietnam. Holding a Master's degree in English Linguistics from Hanoi University, her academic expertise spans Social-Emotional Learning (SEL), AI applications in education, and linguistics. Ms. Huong is particularly passionate about bridging the "well-being gap" in secondary education. Her recent research focuses on the systematic integration of SEL frameworks to enhance student emotional resilience and academic engagement in modern classroom environments.

#29 English-Major Students' Perceptions of AI Tools for Academic Research: A Quantitative Case Study at a Vietnamese University

Tran Thanh Tu[1], Tran Nam Anh[2]

Abstract

Artificial intelligence (AI) tools have entered rapidly into university classrooms, yet little is known about how English-major students use them across the entire chain of academic research activities rather than for writing alone. This study examined perceptions and use of AI tools for academic research among 188 undergraduates at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), Vietnam. A cross-sectional survey, grounded in the Technology Acceptance Model (TAM), self-regulated learning (SRL) theory, and the TPACK framework, measured seven constructs: perceived usefulness, perceived ease of use, teacher's AI integration, self-regulated learning, AI dependence, ethical awareness, and AI usage behavior. Responses were analyzed with descriptive statistics, reliability analysis, and exploratory factor analysis (EFA). Students reported widespread adoption of AI tools, ChatGPT in particular, but used them mainly to search for information and generate ideas rather than to produce finished text. The factor structure suggests that students see technology perceptions, teacher guidance, and self-regulation as connected parts of the same learning experience, rather than as separate factors. The findings extend technology acceptance research to undergraduate research training and show how universities can bring AI into research instruction without weakening academic integrity.

Keywords: artificial intelligence, academic research, technology acceptance model (TAM), English-major students

1. Introduction

1.1. General Context

Artificial intelligence has changed how universities operate and how foreign languages are taught (Arif, 2025). Since its launch in November 2022, ChatGPT reached 100 million users within just two months, making it the fastest-growing consumer application in history (Heaven, 2023). Moreover, modern large language models can do much more than simply retrieve information. Generative AI tools are capable of producing well-structured academic content while offering immediate feedback on grammar and writing style (Kasneci et al., 2023). For EFL students, these capabilities are particularly useful because AI can offer personalized language support, including immediate feedback on writing and opportunities for low-anxiety speaking practice. Nevertheless, growing concerns suggest that dependence on AI may weaken the deep cognitive processing required for authentic language acquisition. The question facing language educators is therefore no longer whether students will use AI, but how they can be guided to use it in ways that enhance rather than replace authentic learning (Najmiah & Al-Anbari, 2025).

Despite the transformative potential of Genarative AI, the pace of adoption has far exceeded the establishment of adequate pedagogical and ethical standards (Alkhatib et al., 2026). UNESCO's Guidance for Generative AI in Education and Research (2023) stressed the need for educational institutions to validate Genarative AI systems on their ethical and

[1] *Department of English, Faculty of Foreign Languages, HUFLIT, Ho Chi Minh City, Vietnam*
[2] *Department of English, Faculty of Foreign Languages, HUFLIT, Ho Chi Minh City, Vietnam.*

pedagogical acceptability, yet noted that fewer than 10% of schools and universities worldwide had formal guidance on AI at the time of publication (Silva & Janes, 2023).

1.2. Recent Studies

1.2.1. International Context

Recent studies suggest that generative AI, particularly ChatGPT, is changing the way EFL students approach academic writing (Tanharat & Phootirat, 2025). Using a mixed-methods approach based on Sociocultural Theory (SCT), the authors investigated 50 senior English-major undergraduates in Thailand, finding that students primarily use ChatGPT as a mediational tool to self-regulate their writing process, specifically for paraphrasing, brainstorming, and enhancing coherence (Tanharat & Phootirat, 2025). However, the authors noted that students maintain a cautious, critical attitude toward ethical concerns such as plagiarism and the reliability of AI-generated information. Ultimately, the authors argued that while AI offers useful cognitive support, human feedback remains essential for developing advanced academic writing proficiency.

Experimental evidence indicates that students who used AI under teacher guidance achieved significantly better learning outcomes compared to those in traditional settings (Guan et al., 2025). Working with 66 high school students in China, the authors found that the experimental group using AI under teacher guidance scored significantly higher in reading, writing, and overall English proficiency than the control group taught in a conventional format. According to the authors, despite the acknowledged benefits of AI in facilitating task completion and generating ideas, the majority of students still preferred the presence of human teachers. These results suggest that teachers remain essential in providing emotional support and pedagogical guidance that AI cannot replace (Guan et al., 2025).

In the context of postgraduate education, Alghamdi & Madini (2026) explored the perceptions of Master of Arts in TESOL students regarding the use of QuillBot for academic writing and publishing. Through semi-structured interviews with 15 female students in Saudi Arabia and applying the Technology Acceptance Model (TAM), the study reveals that students highly appreciate QuillBot for enhancing grammatical accuracy, sentence variety, and paraphrasing skills (Alghamdi & Madini, 2026). Nevertheless, the authors reported that participants had serious reservations about the risk of semantic distortion, academic integrity issues, and the potential erosion of independent writing skills due to overreliance. Consequently, the authors recommended that AI tools should be integrated strictly as supplementary aids rather than primary writing instruments, with careful evaluation from users.

How ChatGPT can support a broad range of language skills was examined by Ahmed et al. (2025) in a mixed-methods study involving 100 students at the Virtual University in Pakistan. According to the authors, interactions with ChatGPT significantly improved students' vocabulary, grammar, writing, and reading comprehension within a short timeframe. Results suggest that AI creates a personalized and low-pressure learning environment with real-time feedback, which is particularly beneficial for language development (Ahmed et al., 2025). However, the authors also cautioned against technical limitations, such as occasional inaccuracies and the inherent absence of authentic human interaction, which remains a vital component of holistic language acquisition.

Exploring the psychological drivers of AI adoption, Wang et al. (2025) applied the Integrated Model of Technology Acceptance (IMTA) to investigate the motivations of 347 university students in China using ChatGPT for EFL learning. Findings indicate that intrinsic motivation plays a dominant role in shaping students' intention to continue using AI tools,

significantly overriding traditional extrinsic factors like perceived usefulness or ease of use. According to the authors, when students experience an engaging, human-like interaction with AI, they are more motivated to adopt it for various academic tasks, including speaking practice and writing enhancement (Wang et al., 2025). Despite this enthusiasm, the authors also noted students' self-awareness regarding the risks of over-dependence and AI hallucinations, advocating for the strategic and guided integration of AI in higher education.

1.2.2. Vietnamese Context

Studies conducted in Vietnam indicate that AI tools offer both significant benefits and notable challenges for EFL students (X. H. N. Thi et al., 2025). Employing a mixed-methods approach with 314 English-majored students, evidence indicates that tools such as ChatGPT and QuillBot are widely used across different stages of academic writing, particularly for complex assignments such as research papers and literature reviews (X. H. N. Thi et al., 2025). According to the authors, there is a positive correlation between the frequency of AI usage and perceived benefits, including improved grammatical accuracy, idea generation, and reduced writing anxiety. However, the authors also cautioned that students face risks such as overreliance, information inaccuracy, and plagiarism.

In a study based on the Technology Acceptance Model (TAM), Vo & Nguyen (2024) explored the perceptions of 369 English-majored students in Central Vietnam regarding ChatGPT. According to the authors, although students exhibited highly positive attitudes and found the tool easy to use, they maintained a largely neutral stance on its overall usefulness, specifically recognizing that AI provided little help for listening and speaking skills (Vo & Nguyen, 2024). In particular, the authors reported that students' class levels significantly influenced their perceptions, whereas gender did not create a significant difference. To adopt ChatGPT effectively, the authors advocated for the development of specific teaching strategies, careful monitoring, and maintaining a critical balance between AI assistance and human interaction in the language learning process.

Applying the Unified Theory of Acceptance and Use of Technology (UTAUT) framework, T. T. M. Thi (2025) investigated AI usage among 53 English majors at a public university in Ho Chi Minh City. According to the author, learners predominantly utilized generative AI for self-study and homework to check grammar and brainstorm ideas, rather than for direct translation (T. T. M. Thi, 2025). The author found high levels of technology acceptance across both male and female students, emphasizing that instructors' recommendations wielded a much more significant influence on students' intentions to use AI compared to peer suggestions. Nevertheless, the author also drew attention to serious concerns in the qualitative data regarding AI-generated misinformation, unnaturally overcomplicated expressions, and the risk of tech-overreliance diminishing learners' critical thinking and creativity.

1.3. Bibliometric Evidence for the Research Gap

Figure 1

Temporal Overlay Visualization of Keyword Co-occurrence in AI Research

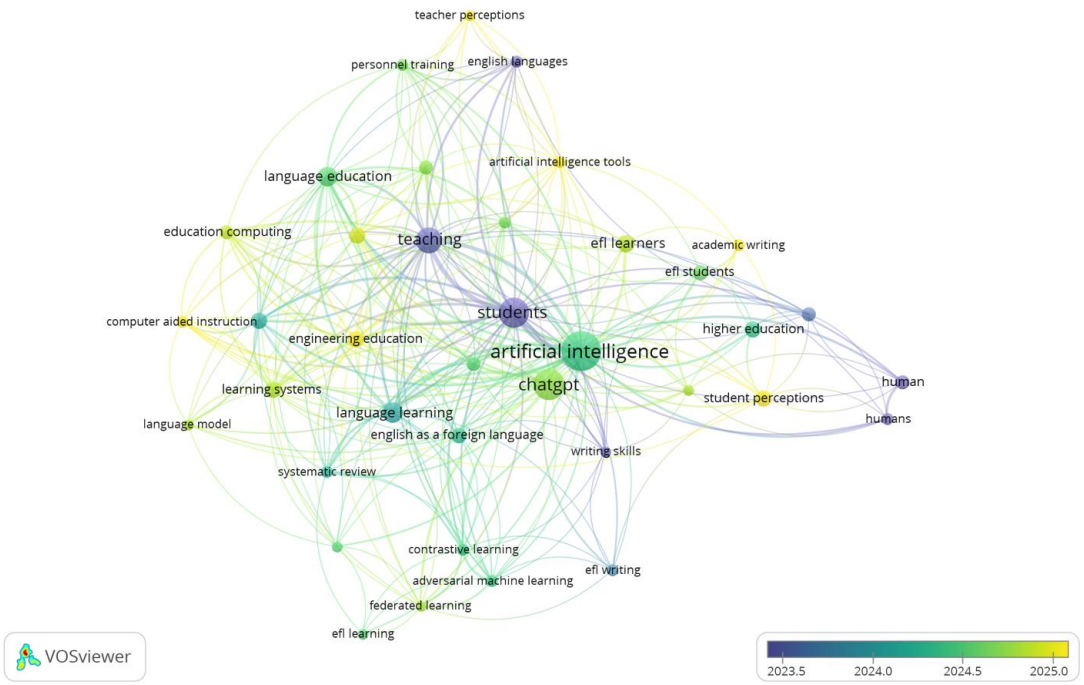


Table 1

Thematic Clusters Identified in the Bibliometric Analysis of AI Research in Language Education

Cluster	Main Theme	Keywords
Cluster 1	AI, ChatGPT & Student Perceptions in Higher Education	academic writing; artificial intelligence; chatgpt; efl learners; efl students; higher education; human; humans; motivation; structural equation modeling; student perceptions
Cluster 2	Language Learning, EFL Writing & Technology Acceptance	adversarial machine learning; contrastive learning; efl learning; efl writing; english as a foreign language; federated learning; language learning; systematic review; technology acceptance; writing skills
Cluster 3	AI in Education, Teaching & Teacher Perspectives	artificial intelligence in education; artificial intelligence tools; empirical studies; english languages; generative artificial intelligence; language education; personnel training; teacher perceptions; teaching

Cluster 4	E-learning Systems & Educational Technology Infrastructure	case studies; computer-aided instruction; e-learning; education computing; engineering education; language model; learning systems; students

The VOSviewer bibliometric analysis indicates that research on artificial intelligence (AI) in language education has developed rapidly and formed several clear research directions. The co-occurrence network identifies four main thematic clusters: (1) AI, ChatGPT, and student perceptions in higher education; (2) language learning, EFL writing, and technology acceptance; (3) AI in education, teaching practices, and teacher perceptions; and (4) e-learning systems and technology-enhanced learning environments.

The overlay visualization provides further insights into the temporal evolution of the field. Earlier research tends to focus on general concepts such as students, teaching, and writing skills, while more recent studies increasingly emphasize academic writing, student perceptions, teacher perceptions, and AI tools. This shift indicates a transition from general exploratory studies toward more focused investigations of user experience, pedagogical implications, and the practical integration of specific AI applications.

1.4. Research Gap and Significance of the Study

First, existing empirical research has largely focused on AI's role in supporting general language skills, particularly writing-related tasks such as grammar correction, paraphrasing, and idea generation (Ahmed et al., 2025; Tanharat & Phootirat, 2025; X. H. N. Thi et al., 2025). However, less attention has been given to how English-major students use AI tools for advanced academic research tasks, including developing research questions, designing methodologies, analyzing data, and synthesizing scholarly literature. This gap is also reflected in the bibliometric co-occurrence network, where keywords such as "academic writing," "writing skills," and "EFL writing" are highly visible, while terms related to research methodology and research skills are largely absent.

Second, the review of existing studies reveals a widespread use of descriptive, survey-based quantitative designs that examine general attitudes toward AI tools (Alharbi, 2024; Najmiah & Al-Anbari, 2025). In the Vietnamese context specifically, studies by Vo & Nguyen (2024) have applied established acceptance frameworks such as TAM and UTAUT to measure general perceptions. However, these studies have not fully considered pedagogical factors and students' self-regulation in AI-supported learning (Vo & Nguyen, 2024).

Third, Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT) presents a particularly significant site for such an investigation. As a university where English-major students are concurrently trained in linguistic competence and exposed to technology-enhanced learning environments, HUFLIT represents a demographic that heavily utilizes AI tools while being directly impacted by their influence on academic writing, linguistic feedback, and research practices.

1.5. Research Objectives and Questions

The first objective is to examine how English-major students at HUFLIT perceive the benefits and challenges of AI-supported academic research (Bui & Tong, 2025). The second objective is to evaluate the psychometric properties and underlying factor structure of the seven-construct measurement instrument. The third objective is to derive theoretically grounded and practically actionable implications from the empirical findings for the ethical

and effective integration of AI tools in English language education within Vietnamese higher education institutions.

To operationalize these objectives, the study addresses the following research questions:

RQ1: How do English-major students at HUFLIT perceive the benefits and challenges of AI-supported academic research, as measured through the constructs of Perceived Usefulness, Perceived Ease of Use, Teacher's AI Integration, Self-Regulated Learning, AI Dependence, Ethical Awareness, and AI Usage Behavior?

RQ2: To what extent do the seven theoretically hypothesized constructs demonstrate internal consistency and empirical distinctiveness, and what underlying factor structure emerges from exploratory factor analysis of the data?

RQ3: What are the theoretical and practical implications of the empirical findings for integrating AI tools ethically and effectively in English language education at Vietnamese higher education institutions?

2. Literature Review

2.1. The Paradigm Shift in AI-Mediated Language Education

In the domain of second language writing, generative AI tools have expanded their influence far beyond surface-level grammar correction to affect the conceptual content and structural organization of student work (Najmiah & Al-Anbari, 2025). generative AI has necessitated a fundamental redesign of how instructors approach the teaching and assessment of writing skills, as students now have access to tools that can produce polished academic prose with minimal human input (Kohnke et al., 2023). Beyond written production, learners increasingly perceive AI as a transformative tool for oral proficiency. For instance, studies in Vietnam indicate that students utilize generative AI platforms to improve their speaking skills through interactive, low-stakes practice that reduces performance anxiety and provides opportunities for sustained conversational engagement outside formal classroom settings.

2.2. Theoretical Perspectives on Technology Acceptance

2.2.1. Core Determinants within the TAM Framework

The Technology Acceptance Model (TAM), established by Davis (1989), remains the primary theoretical lens for evaluating user adoption of educational technologies. The model posits that two core constructs: perceived usefulness (PU) and perceived ease of use (PEOU), serve as the essential determinants of a user's behavioral intention to engage with new technology (Venkatesh & Davis, 2000). In the context of EFL education, perceived usefulness is typically operationalized as the belief that AI tools directly enhance linguistic proficiency, academic efficiency, and research productivity (Vo & Nguyen, 2024).

2.2.2. Expanding TAM in the Generative AI Era

In the context of ChatGPT adoption in higher education, studies utilizing the Unified Theory of Acceptance and Use of Technology (UTAUT) have shown that alongside performance expectancy, students' intention to use is significantly influenced by facilitating conditions and social influence. Furthermore, result demonstrability has been identified as an important predictor of learners' usage behavior. Additionally, students' decisions to accept AI are strongly driven by emotional factors as well as concerns about ethics, privacy, and academic integrity (Haetami & Aghata, 2025; Kohnke et al., 2023; Ravšelj et al., 2025).

2.3. AI Support and the Risk of Shortcut Learning in Self-Regulated Learning

Self-regulated learning (SRL), defined by Zimmerman (2002) as the self-initiated process through which learners set goals, monitor progress, and evaluate outcomes, has emerged as a critical construct for understanding how students engage with AI tools in educational settings (Zimmerman, 2002).

Generative AI offers significant potential as a scaffolding mechanism that supports the cyclical SRL process of forethought, performance, and self-reflection (Kasneci et al., 2023). AI-powered tools can provide personalized feedback, real-time progress monitoring, and metacognitive prompts that facilitate learners' self-regulatory development. For many English-major students in Vietnam, ChatGPT serves as a constant learning companion that supports autonomy through immediate writing feedback, research planning support, and self-directed vocabulary and grammar practice (Vo & Nguyen, 2024). Consistent with this view, recent systematic reviews indicate that Generative AI has strong potential to enhance SRL by helping students with goal-setting, strategy selection, and reflective evaluation across different academic contexts (Chiu, 2024; Kien, 2025; Ravšelj et al., 2025).

2.4. Ethical Integrity and Policy Realities

International frameworks from UNESCO (2023) and the OECD (2023) advocate for a human-centered and ethical approach to AI in education, emphasizing principles of transparency, accountability, inclusivity, and the protection of learner agency (OECD, 2023; UNESCO, 2023). However, a gap remains between these policy-level principles and classroom realities, particularly in developing countries where institutional resources, teacher training, and digital infrastructure may be insufficient. In Vietnamese higher education specifically, students frequently use AI tools, but their use is often not guided by clear institutional policies. While students recognize the utility of Generative AI for academic tasks, they also confront concerns over academic integrity, the reliability of AI-generated information, and the absence of clear boundaries regarding acceptable use (X. H. N. Thi et al., 2025).

3. Methodology

3.1. Research Design

A quantitative cross-sectional design was employed to examine students' perceptions of AI tools in academic research at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT).

3.2. Participants and Sampling

The study involved 188 undergraduate English-major students at HUFLIT. Participants were selected using a convenience sampling approach, focusing on students with prior exposure to both language learning and research activities.

3.3. Instrument Development

The survey instrument consisted of two sections. The first section collected demographic information, including gender, year of study, prior experience with AI tools, frequency of AI use, most frequently used AI tool, and purposes of AI usage in academic research (multiple-response items). The second section comprised 27 Likert-scale items measured on a five-point scale (1 = Strongly Disagree to 5 = Strongly Agree), distributed across seven theoretically grounded constructs: Perceived Usefulness (PU, 4 items), Perceived Ease of Use (PEOU, 4 items), Teacher's AI Integration (TPACK, 5 items), Self-Regulated Learning

(SRL, 4 items), AI Dependence (AD, 3 items), Ethical Awareness (ETH, 3 items), and AI Usage Behavior (AUB, 4 items).

3.4. Data Collection Procedure

Data were collected during the second semester of the 2025–2026 academic year through paper survey and Google Forms.

3.5. Data Analysis

Data analysis was conducted using IBM SPSS Statistics (Version 26) and proceeded through three sequential phases, each corresponding to a specific research question.

4. Results

4.1. Descriptive Statistics

4.1.1. Demographic Characteristics of Respondents

The sample consisted of 188 undergraduate students. The sample was predominantly female (n = 140, 74.5%), with male participants comprising 25.5% (n = 48). The majority of respondents were third-year students (n = 158, 84.0%), while fourth-year students accounted for 16.0% (n = 30). A considerable proportion of respondents (96.8%, n = 182) reported having prior experience with AI tools for research purposes. In terms of frequency of use, 45.7% (n = 86) indicated using AI "often," followed by 26.6% (n = 50) who used it "very often" and 23.4% (n = 44) who used it "sometimes." ChatGPT was the most frequently used AI tool (70.2%, n = 132), followed by Grammarly (11.7%, n = 22), Google Translate (9.0%, n = 17), and Perplexity (9.0%, n = 17).

Table 2

Demographic Characteristics of Respondents (N = 188)

Variable	Category	n	%	Cum. %
Gender	Male	48	25.5	25.5
	Female	140	74.5	100.0
Year of Study	Third year	158	84.0	84.0
	Fourth year	30	16.0	100.0
AI Usage Experience	Yes	182	96.8	96.8
	No	6	3.2	100.0
Frequency of AI Use	Never	4	2.1	2.1
	Rarely	4	2.1	4.2
	Sometimes	44	23.4	27.6
	Often	86	45.7	73.3
	Very Often	50	26.6	99.9
Most Used AI Tool	ChatGPT	132	70.2	70.2
	Grammarly	22	11.7	81.9
	Google Translate	17	9.0	90.9
	Perplexity	17	9.0	99.9

4.1.2. Purposes of AI Usage in Scientific Research

Respondents reported multiple purposes for using AI in scientific research activities. The most prevalent purpose was searching for information (77.1%), followed by generating ideas and topics (72.9%), and translating academic content (61.7%). Writing or revising text and summarizing articles were reported by approximately 55% of respondents, while creating references or citations was the least common purpose (45.7%).

Table 3

Purposes of AI Usage in Scientific Research (N = 188, Multiple Responses)

Purpose	n (Yes)	% (Yes)	% (No)
Searching for information	145	77.1	22.9
Generating ideas / topics	137	72.9	27.1
Translating academic content	116	61.7	38.3
Writing or revising text	105	55.9	44.1
Summarizing articles	104	55.3	44.7
Creating references / citations	86	45.7	54.3

4.1.3. Descriptive Statistics of Likert-Scale Items

Table 4 presents the descriptive statistics for all 27 Likert-scale items across seven constructs. Overall, the mean scores ranged from 3.67 (AD1) to 4.26 (PU1), indicating that respondents generally expressed agreement to strong agreement with most survey items. The standard deviations ranged from 0.668 to 1.049, suggesting moderate variability in responses. All items exhibited negative skewness values, indicating a tendency toward higher scores.

Table 4

Descriptive Statistics of Likert-Scale Items (N = 188)

Item	M	SD	Skewness	Kurtosis	Min	Max
PU1	4.26	0.839	-1.280	2.232	1	5
PU2	4.24	0.823	-1.005	1.151	1	5
PU3	4.18	0.881	-0.968	0.742	1	5
PU4	4.22	0.878	-1.209	1.895	1	5
PEOU1	4.20	0.782	-0.915	1.429	1	5
PEOU2	4.19	0.757	-0.335	-1.185	3	5
PEOU3	4.04	0.892	-0.713	0.248	1	5
PEOU4	4.02	0.942	-0.741	0.201	1	5
TPACK1	4.05	0.861	-0.398	-0.614	1	5
TPACK2	4.05	0.876	-0.683	0.250	1	5
TPACK3	4.08	0.858	-0.667	0.273	1	5
TPACK4	4.10	0.870	-0.725	0.374	1	5
TPACK5	4.16	0.814	-0.662	0.010	1	5
SRL1	4.13	0.668	-0.258	-0.159	2	5

SRL2	4.09	0.801	-0.601	0.103	1	5
SRL3	4.16	0.841	-0.855	0.684	1	5
SRL4	4.14	0.852	-0.834	0.544	1	5
AD1	3.67	1.049	-0.849	5.515	1	5
AD2	3.93	0.946	-0.763	0.249	1	5
AD3	4.01	0.933	-0.894	0.616	1	5
ETH1	4.16	0.871	-0.866	0.413	1	5
ETH2	4.07	0.889	-0.651	-0.198	1	5
ETH3	4.10	0.897	-0.854	0.422	1	5
AUB1	4.12	0.873	-0.839	0.543	1	5
AUB2	4.13	0.847	-0.740	0.121	1	5
AUB3	4.11	0.874	-0.825	0.470	1	5
AUB4	4.01	0.903	-0.637	-0.020	1	5

One exception should be acknowledged. Item AD1 showed a kurtosis value of 5.515, exceeding the ±2 range commonly expected under strict univariate normality. This indicates that most responses were concentrated around the agreement categories, while a small number of students selected the lowest response options, creating a more peaked distribution with heavier tails. Such a pattern can occur in Likert-scale items that address sensitive self-perceptions, such as dependence (Blanca et al., 2013). Since the value was still below the |kurtosis| < 7 benchmark suggested by West et al. (1995), and because reliability analysis and exploratory factor analysis are relatively robust to moderate non-normality at this sample size, AD1 was retained without transformation.

4.2. Reliability Analysis

Cronbach's alpha coefficients ranged from .852 (AI Dependence) to .935 (Teacher's AI Integration), indicating good to excellent internal consistency across all seven constructs. All constructs met the established reliability thresholds (Nunnally, 1978). Moreover, all corrected item-total correlations (CITC) exceeded .30, and no item deletion would meaningfully improve the alpha values. These results support the retention of all items for subsequent analysis.

Table 5

Summary of Reliability Analysis by Construct

Construct	Items	Cronbach's α	CITC Range	Assessment
Perceived Usefulness (PU)	4	0.904	.757–.797	Excellent
Perceived Ease of Use (PEOU)	4	0.879	.723–.753	Good
Teacher's AI Integration (TPACK)	5	0.935	.814–.858	Excellent
Self-Regulated Learning (SRL)	4	0.892	.711–.790	Good
AI Dependence (AD)	3	0.852	.707–.795	Good
Ethical Awareness (ETH)	3	0.912	.800–.841	Excellent
AI Usage Behavior (AUB)	4	0.922	.777–.843	Excellent

Note. CITC = Corrected Item-Total Correlation. All CITC values > .30. Internal consistency thresholds follow the standard .70 reliability benchmark (Nunnally, 1978).

Table 6

Item-Level Reliability Statistics

Construct	Item	CITC	α if Deleted	Decision
Perceived Usefulness (PU)	PU1	0.797	0.872	Retain
	PU2	0.792	0.874	Retain
	PU3	0.757	0.886	Retain
	PU4	0.795	0.873	Retain
Perceived Ease of Use (PEOU)	PEOU1	0.753	0.840	Retain
	PEOU2	0.745	0.845	Retain
	PEOU3	0.752	0.839	Retain
	PEOU4	0.723	0.854	Retain
Teacher's AI Integration (TPACK)	TPACK1	0.814	0.923	Retain
	TPACK2	0.838	0.918	Retain
	TPACK3	0.819	0.922	Retain
	TPACK4	0.858	0.915	Retain
	TPACK5	0.814	0.923	Retain
Self-Regulated Learning (SRL)	SRL1	0.711	0.879	Retain
	SRL2	0.760	0.862	Retain
	SRL3	0.788	0.851	Retain
	SRL4	0.790	0.850	Retain
AI Dependence (AD)	AD1	0.707	0.847	Retain
	AD2	0.795	0.723	Retain
	AD3	0.748	0.784	Retain
Ethical Awareness (ETH)	ETH1	0.800	0.899	Retain
	ETH2	0.841	0.863	Retain
	ETH3	0.839	0.861	Retain
AI Usage Behavior (AUB)	AUB1	0.832	0.894	Retain
	AUB2	0.843	0.891	Retain
	AUB3	0.829	0.895	Retain
	AUB4	0.777	0.913	Retain

Note. CITC = Corrected Item-Total Correlation; α if Deleted = Cronbach's Alpha if Item Deleted. All items retained based on CITC > .30 criterion.

4.3. Exploratory Factor Analysis (EFA)

4.3.1. Sampling Adequacy and Sphericity

Prior to exploratory factor analysis, the suitability of the data was examined. The Kaiser-Meyer-Olkin (KMO) measure was .917, indicating high sampling adequacy. Bartlett's Test of Sphericity was statistically significant ($\chi^2 = 4672.900$, df = 351, $p < .001$), suggesting that the correlation matrix was appropriate for factor analysis.

Table 7

KMO and Bartlett's Test

Test	Statistic	Value
Kaiser-Meyer-Olkin Measure	—	.917
Bartlett's Test of Sphericity	Approx. χ^2	4672.900
	df	351
	Sig.	< .001

4.3.2. Total Variance Explained

Principal Component Analysis (PCA) with Varimax rotation was employed for factor extraction. Based on the eigenvalue-greater-than-one criterion, five components were extracted, collectively accounting for 75.127% of the total variance. The first component explained the largest proportion of variance (50.255% before rotation; 27.129% after rotation). After Varimax rotation, the variance was more evenly distributed across the five factors. The rotation converged in five iterations.

Table 8

Total Variance Explained

Comp.	Eigenvalue	% Variance	Cum. %	Rot. Total	Rot. % Var.	Rot. Cum. %
1	13.569	50.255	50.255	7.325	27.129	27.129
2	2.433	9.012	59.268	3.871	14.335	41.464
3	1.628	6.031	65.299	3.764	13.942	55.406
4	1.444	5.349	70.648	2.873	10.641	66.047
5	1.209	4.479	75.127	2.452	9.080	75.127

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Only components with eigenvalues ≥ 1.0 are shown.

4.3.3. Rotated Component Matrix

Table 9 displays the rotated component matrix using Varimax rotation with Kaiser normalization. Factor loadings below .40 are not reported for clarity. The five extracted factors demonstrated a clear structure. Factor 1 consisted of 13 items from PU, TPACK, and SRL, with loadings ranging from .543 to .826. Factor 2 included four PEOU items (.697–.811), Factor 3 included four AUB items (.800–.814), Factor 4 included three ETH items (.787–.864), and Factor 5 included three AD items (.864–.913). All items loaded clearly on their intended factors, with no problematic cross-loadings.

Table 9

Rotated Component Matrix (Varimax Rotation)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
SRL4	0.826				
SRL2	0.778				
PU1	0.775				
SRL3	0.755				
TPACK5	0.727				
TPACK4	0.700				
PU3	0.690				
PU2	0.678				
SRL1	0.652				
TPACK3	0.630				
PU4	0.626				
TPACK2	0.572				
TPACK1	0.543				
PEOU3		0.811			
PEOU1		0.716			
PEOU2		0.709			
PEOU4		0.697			
AUB2			0.814		
AUB4			0.807		
AUB3			0.802		
AUB1			0.800		
ETH2				0.864	
ETH3				0.809	
ETH1				0.787	
AD2					0.913
AD3					0.873
AD1					0.864

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Factor loadings < .40 are suppressed. Rotation converged in 5 iterations.

4.4. Summary of Key Findings

The quantitative analyses revealed several key findings. First, the descriptive statistics indicated that respondents generally viewed AI tools positively in the context of academic research. Mean scores for most Likert-scale items were consistently above 4.00, except for AI Dependence, which ranged from 3.67 to 4.01. Second, the reliability analysis confirmed strong internal consistency across all seven constructs, with Cronbach's alpha values ranging from .852 to .935. Third, the exploratory factor analysis extracted five factors, which together

explained 75.127% of the total variance. Although the study initially proposed a seven-construct structure, the empirical results reduced this structure to five factors. Factor 1 combined items from Perceived Usefulness, Teacher's AI Integration, and Self-Regulated Learning; Factor 2 retained all Perceived Ease of Use items; Factor 3 comprised all AI Usage Behavior items; Factor 4 grouped all Ethical Awareness items; and Factor 5 included all AI Dependence items.

5. Discussion

5.1. Student Perceptions of AI in Academic Research

The very high AI adoption rate among HUFLIT English-major students (96.8%) and their frequent use of AI are consistent with global trends, where AI adoption among university students has been reported at 80–92%. Students mainly used AI for information searching (77.1%) and idea generation (72.9%), rather than for producing complete texts. This suggests that AI functions primarily as a cognitive support tool during the exploratory stages of research, rather than as a tool for automating final outputs. Mean scores were above 4.00 across all constructs except AI Dependence (M = 3.67–4.01), indicating positive perceptions of AI, together with awareness of the risks of excessive reliance. This finding reflects the scaffolding–shortcut paradox identified in the literature (Alghamdi & Madini, 2026; T. T. M. Thi, 2025; Wang et al., 2025).

5.2. Factor Structure and Integrated AI Research Competence

The most important finding is the five-factor EFA structure departing from the hypothesized seven-construct model. While PEOU, AUB, Ethical Awareness, and AI Dependence were cleanly reproduced, PU, TPACK, and SRL converged into a single composite factor, Integrated AI Research Competence, suggesting students experience technology perceptions, pedagogical guidance, and metacognitive regulation as interrelated facets of a unified competence rather than separate dimensions. The clean separation of Ethical Awareness (loadings .787–.864) and AI Dependence (loadings .864–.913) as distinct factors suggests that ethical cognition and dependency perceptions may operate through different psychological mechanisms.

5.3. Implications

Theoretically, the Integrated AI Research Competence construct extends TAM by showing that PU, pedagogical scaffolding, and SRL may collapse into a single latent dimension when instructional integration is sufficiently strong. Future research should employ CFA and SEM to validate this construct across contexts.

Practically, AI competence programs should adopt holistic approaches developing technology perceptions, instructional guidance, and self-regulation concurrently through scaffolded research projects mirroring Zimmerman's (2002) SRL cycle. AI ethics should be embedded as a dedicated curricular component, not treated incidentally. Students should adopt deliberate strategies for evaluating AI outputs and periodically completing tasks without AI to mitigate dependency risks.

5.4. Limitations

Several limitations should be acknowledged. First, the study relied on a convenience sample from a single institution (HUFLIT), which limits the generalizability of the findings. Second, the use of self-report measures introduces the possibility of social desirability bias, particularly on sensitive items related to AI dependence and ethical awareness.

6. Conclusion

This study examined 188 undergraduate students' perceptions and use of AI tools in scientific research through a quantitative survey grounded in the TAM, self-regulated learning theory, and the TPACK framework across seven constructs. In general, the findings indicate that AI tools, particularly ChatGPT, have become widely integrated into students' research workflows (96.8% usage rate), mainly for information searching, idea generation, and translation, positioning AI as an augmentative rather than replacement resource. The measurement instrument showed strong reliability, with Cronbach's alpha values ranging from .852 to .935. While exploratory factor analysis produced a five-factor structure explaining 75.127% of total variance, departing from the hypothesized seven-construct model. The most important difference was that Perceived Usefulness, Teacher's AI Integration, and Self-Regulated Learning converged into a single composite factor termed Integrated AI Research Competence. This suggests that students see AI usefulness, teacher guidance, and self-regulation as closely connected parts of one broader competence.

Theoretically, this study extends the TAM to AI-assisted undergraduate research, introduces Integrated AI Research Competence as a novel composite construct, and provides evidence that ethical awareness and AI dependence function as empirically distinct factors, which may operate through different psychological mechanisms. Practically, the findings suggest that AI competence training should adopt a holistic approach simultaneously addressing perceived value, instructional scaffolding, and self-regulation rather than treating these dimensions in isolation. AI ethics should be integrated into formal research methodology courses instead of being presented as an additional topic. Students should also be encouraged to maintain critical autonomy by evaluating AI outputs carefully, checking them against primary sources, and occasionally completing tasks without AI assistance to preserve independent research skills.

References

Ahmed, K., Akhter, S., Siwok, S. R., Tambi, F., Khan, A. A., & Farooqi, U. u. R. (2025). Role of ChatGPT in enhancing English language learning: A case study of university students. *Forum for Linguistic Studies, 7*(7), 460–469. https://doi.org/10.30564/fls.v7i7.9165

Alghamdi, A., & Madini, A. A. (2026). QuillBot as an AI writing support tool in English academic writing: MA TESOL students' perceived challenges and solutions. *Arab World English Journal*, Special Issue on Artificial Intelligence, (3), 303–315. https://doi.org/10.24093/awej/AI3.19

Alharbi, M. (2024). The role of artificial intelligence in advancing English as a foreign language teaching at Saudi universities. *World Journal on Educational Technology: Current Issues, 16*(3), 181–200. https://doi.org/10.18844/wjet.v16i3.9311

Alkhatib, O. J., Shafiq, M. H., Hameed, S., & Khan, A. M. (2026). The dark side of AI in education: Unintended consequences on English language learning and teaching. *Journal of Applied Linguistics and TESOL, 9*(1), 1–14. https://doi.org/10.63878/jalt1782

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology, 9*(2), 78–84. https://doi.org/10.1027/1614-2241/a000057

Bui, T. T. U., & Tong, T. V. A. (2025). The impact of AI writing tools on academic integrity: Unveiling English-majored students' perceptions and practical solutions. *AsiaCALL Online Journal, 16*(1), 83–110. https://doi.org/10.54855/acoj.251615

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence, 6*, Article 100197. https://doi.org/10.1016/j.caeai.2023.100197

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Guan, L., Lee, J. C. K., Zhang, Y., & Gu, M. M. (2025). Investigating the tripartite interaction among teachers, students, and generative AI in EFL education: A mixed-methods study. *Computers and Education: Artificial Intelligence, 8*, Article 100384. https://doi.org/10.1016/j.caeai.2025.100384

Haetami, A., & Aghata, F. (2025). Ethical and data security analysis in the implementation of generative AI in higher education environments. *MSJ: Majority Science Journal, 3*(4), 177–186. https://doi.org/10.61942/msj.v3i4.482

Heaven, W. D. (2023, March 3). The inside story of how ChatGPT was built from the people who made it. *MIT Technology Review*. https://www.technologyreview.com/2023/03/03/1069311/inside-story-oral-history-how-chatgpt-built-openai/

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models

for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kien, T. T. (2025). The impact of artificial intelligence on the English self-study process at Nguyen Tat Thanh University. *Journal of Science and Technology, 8*(5), 52–59. https://doi.org/10.55401/hxn2hh66

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Najmiah, L., & Al-Anbari, A. H. J. (2025). Narrative explorations of EFL learners' engagement with AI tools in developing writing proficiency. *DUTIES: Education and Humanities International Journal, 1*(2), 42–62. https://doi.org/10.70152/duties.v1i2.222

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

OECD. (2023). *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. https://doi.org/10.1787/c74f03de-en

Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., A Iahad, N., Abdulla, A. A., Akopyan, A., Aldana Segura, M. W., AlHumaid, J., Allam, M. F., Alló, M., Andoh, R. P. K., Andronic, O., Arthur, Y. D., Aydın, F., Badran, A., Balbontín-Alvarado, R., Ben Saad, H., … Aristovnik, A. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE, 20*(2), Article e0315011. https://doi.org/10.1371/journal.pone.0315011

Silva, A. de O., & Janes, D. dos S. (2023). Artificial intelligence in education: What are the opportunities and challenges? *Review of Artificial Intelligence in Education, 5*(00), e018. https://doi.org/10.37497/rev.artif.intell.educ.v5i00.18

Tanharat, N., & Phootirat, P. (2025). Integrating generative AI in EFL academic writing: Thai English-major students' purposes, perceptions, and experiences with ChatGPT. *rEFLections, 32*(3), 1891–1916. https://doi.org/10.61508/refl.v32i3.285982

Thi, T. T. M. (2025). Understanding Vietnamese English majors' use and acceptance of AI-powered tools for English academic writing at university. *Taiwan International ESP Journal, 15*(2), 29–54. https://doi.org/10.6706/TIESPJ.202412_15(2).0002

Thi, X. H. N., Thien, H. V. H., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through AI-powered tools: Perceived benefits and challenges among Vietnamese EFL students. *Discover Education, 4*(1), Article 472. https://doi.org/10.1007/s44217-025-00905-9

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vo, A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/fr1rkj58

Wang, L., Zou, B., Wang, C., & Huang, F. (2025). Understanding learner motivation and intention to use ChatGPT for English language learning. *Discover Computing, 28*(1), Article 293. https://doi.org/10.1007/s10791-025-09805-w

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

#34 University Students' Attitude and Intentions Towards The Use of Chatgpt and Canva in a Project-based Learning Activity

Vi Le Bao[1]

Abstract

There is a growing interest in applications of Artificial Intelligence technology (AI) in education, especially in the English as a Foreign Language (EFL) environment, where user willingness and intentions play a significant role in successful AI integration. This study aimed to understand the attitudes and behavioral intentions of Business English students at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Vietnam, regarding their use of AI tools (e.g., ChatGPT and Canva) in a project-based learning activity (PBL). 136 Business English students took part in a Likert scale survey which was constructed from the extended Technology Acceptance Model (TAM). The collected data was analyzed using Partial Least Square-Structural Equation Modelling (PLS-SEM) to assess the effects of self-efficacy (SE), creativity support tools (CSTs), perceived usefulness (PU), and perceived ease of use (PEOU) on students' attitudes (Att) and their behavioral intention (BI) toward future AI integration. The results found that Business English students demonstrated both positive attitudes and positive behavioral intentions towards AI technologies, which further highlighted the roles of self-efficacy and creativity support. The study also provided some implications and recommendations for adopting ChatGPT and Canva in higher education.

***Keywords*:** ChatGPT, Canva, Project-based Learning (PBL), Technology Acceptance Model (TAM), attitude, behavioral intention

1. Introduction

Artificial intelligence (AI) has gained public attention since its introduction in the 1950s. AI is defined as the capacity of several related technologies such as digital computers, or computer-controlled robots to execute tasks typically requiring human intelligence (O'Regan, 2016). The applications of AI technologies in education have been identified in various functions, including personalized learning, automated assessment, and administrative tasks (Eager & Brunton, 2023). This integration does not only present technological changes but also signifies a fundamental transformation in the dissemination, evaluation, and use of knowledge (O'Dea & O'Dea, 2023). The literature identifies at least ten ways in applying AI technology in education: (1) automatic grading system, (2) interval reminder, (3) teacher's feedback, (4) virtual teachers, (5) personalized learning, (6) adaptive learning, (7) augmented reality/virtual reality, (8) accurate reading, (9) intelligent campus and (10) distance learning (Yufeia et al., 2020). This transition has created opportunities for increased creativity and efficiency, fostering a more dynamic and interactive learning environment, while also posing challenges for educational institutions worldwide.

The recent release of ChatGPT (Generative Pre-trained Transformer) in 2022 by Open AI has accelerated digitalization and driven advancements in Generative Artificial Intelligence (GAI). According to Miller (2023), GAI comprises AI systems and algorithms which can synthesize realistic outputs and create new and novel contents such as text, images, audio,

[1] *Business English, Foreign Languges, Ho Chi Minh City University of Foreign Languages - Information Technology, Ho Chi Minh City, Vietnam; Email: vilb@huflit.edu.vn*

video, or code, in response to a designated prompt. This helps people to prototype ideas, generate content, or simulate interactive experiences. If ChatGPT, especially version 3.5, is classified as a text producing output of GAI, then image-producing outputs such as Canva, Bing Image Creator, DALL-E are examples of description-based imagery creation. Another significant type of GAI output concerns video production platforms like Runway or Pika Labs (Pika). These technologies can generate motion graphics from framed sequential images; however, they have demonstrated some limitations in producing smooth and coherent video content with free of charge access (Browning-Samoni, 2024).

In English as Foreign Language (EFL) educational contexts, AI tools or AI-driven platforms like ChatGPT or Canva have been extensively used by both educators and students. Recent studies primarily focus on the students' perspectives towards the use of AI tools in acquiring English language skills, including writing, speaking, and reading (Wong et al., 2013; Xiao & Zhi, 2023), whereas other research investigates EFL educators' attitudes toward the application of AI in their lesson preparation and assessments, such as developing lesson plans, creating multimedia presentations, and producing quizzes, or tests (Eager & Brunton, 2023) Notwithstanding its significance, there exists a noticeable lack of studies examining students' attitudes regarding the incorporation of such technologies in specialized university disciplines and in real-world academic tasks. Likewise, it had been observed that students in the author's Selling Skills course had been using tools like ChatGPT 3.5 and Canva to brainstorm and create product marketing postings on social media platforms (i.e. Facebook, Instagram, etc.) during their class projects. To be specific, groups of 5 or 6 students worked together to promote a product of their choice. They had to go through a few steps, including generating ideas, selecting suitable content, visualizing attractive images, designing compelling marketing banners or product templates. These marketing materials were then shared on social media platforms like Facebook or Instagram to advertise their products and to engage with potential customers. This hands-on project was part of the coursework assessment to learn how Business English students apply their knowledge and marketing theory into real-life practices. Acknowledging how these AI technologies could boost student creativity, support personalized learning experiences, and improve their study effectiveness, the author decided to explore Business English majored student' attitudes and intentions on the usage of AI tools in a project-based learning (PBL) activity.

Undoubtedly, a large number of theoretical frameworks have been proposed to unpack individuals' attitudes and intention to adopt a particular technology. Some of the most notable examples include the Technology Acceptance Model (TAM) (Davis, 1989), the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980), and the Theory of Planned Behavior (TPB) (Ajzen, 1985). Among them, the Technology Acceptance Model (TAM) by Davis (1989) is widely recognized as a fundamental framework for understanding technology adoption (Wong et al., 2013). It has been extensively applied in various fields, contexts, and technological tools (Dishaw & Strong, 1999; Venkatesh & Bala, 2008; Wong et al., 2013; Shahzad et al.,2024). In this study, the Technology Acceptance Model (TAM) was utilized to gain an understanding of students' attitude towards the utilization of ChatGPT 3.5 and Canva, while also seeking for opportunities for their future application in PBL activities within the Vietnamese EFL context, such as social media-based product promotion. The research questions guiding this study are as follows:

1. What are students' attitudes toward the use of ChatGPT and Canva in PBL activities?
2. Are they going to use these tools for similar projects in the future?

2. Literature Review and Model Development

2.1. Project-based learning and AI incorporation

Project-based learning (PBL) represents a significant pedagogical shift in educational transformation, positioning students as active learners in their leaning journey. Contrast to traditional lecture-based, teacher-centered, and textbook-driven instructions, PBL enables students to apply their skills and knowledge, collaborate in groups, and synthesize various subject areas to address authentic, real-world problems (Campbell, 2014). Project-Based Learning (PBL) was grounded in constructivist theory by Hmelo-Silver (2004) and was initially introduced by John Dewey in the late 1890s (Simpson & Stack, 2010).The core of PBL lies in its emphasis on active and experiential learning, which requires students to form critical-thinking, self-reliance, responsibility, and discipline through the process of planning, organizing, managing their time, and self-motivating their progress (Thomas, 2000). Project-Based Learning (PBL) allows students to explore topics and questions independently, particularly during the language acquisition process. This approach also generates intentional opportunities for students to practice using the language and to focus on important language aspects like grammar, vocabulary, skills, and learning strategies (Stoller, 2006).

Given the development of AI technology and the rapid growth of the application fields, the combination of artificial intelligence (AI) technologies and project-based learning (PBL) is indeed an inevitable trend and a significant pedagogical approach (Tang et al., 2024). The incorporation of AI enables creativity and exploration for design, problem-solving, decision making, and investigation activities. The use of AI and machine learning (ML) in e learning platforms enhances student interest and retention. These technologies provide personalized feedback and tailored content, both of which are essential determinants to support active learning in project-based contexts (Gligorea et al. 2023). In addition, AI-driven tools such as ChatGPT enable learners to define their own learning outcomes and project topics, thereby fostering autonomy and personalized learning pathways. (Liang& Luo, 2024).

The potential of AI encompasses the development of EFL materials and complementary activities.In an AI-driven digital storytelling strategy, Belda-Medina and Goddard (2024) identified Canva.ai, Fliki, Invideo AI, Animoto, and Jasper A as useful tools for generating digital stories, videos, audio, and vocabulary worksheets. Meanwhile, Twee, Magic School, and Curipod were recommended to produce tasks, exercises, and assignments. These activities, in turn, support leaners' language development, creativity, and autonomy which reflects users' positive attitudes towards AI and its educational benefits in interactive, project-based tasks. Besides, collaborative technologies and AI- driven platforms provide ongoing assistance to students in their projects with peers and educators even beyond the confines of physical classrooms. These technologies not only enhance learners' productivity but also support teachers in tracking and monitoring students'progress more readily, hence providing intime feedback (Cheung & Vogel, 2013). Despite the benefits, it is important to acknowledge that the successful implementation of project work facilitated by digital technologies relies not only on technological access but also on users' digital literacy and their readiness for technology adoption, which remains as critical factors for effective implementation (Belda-Medina, 2021).

2.2 Chat GPT and Canva in EFL education

In EFL settings, ChatGPT presents many opportunities for improving language instructions and learning experiences. As mentioned by Yufeia et al. (2020), tools powered by AI like ChatGPT are effective learning companions, peer tutors, and virtual instructors. They can assist with reading and writing practices such as vocabulary enrichment, grammar check, and idea generation. ChatGPT provides personalized and adaptive feedback on text revision,

essay composition, content recommendation, and guidance on academic paper sections (Xiao & Zhi, 2023). For speaking, ChatGPT creates interactive environments where learners can build their confidence in expression, ease in discussions, and improvement in pronunciation (Monika & Suganthan, 2024). For PBL activities, Liang and Luo (2024) suggested using ChatGPT for business projects like "Exploring the Current Situation and Impact of Cashless Payments in Hong Kong". Students used ChatGPT to generate ideas and to formulate questionnaires for interviewing merchants and consumers. This was all meant toward fostering the development of language communication skills needed in professional environments, while concurrently building business-related vocabulary specific to the business world. In a similar manner, ChatGPT consists of a number of features that optimize the design process for projects like marketing promotional posts. Learners can come up with design ideas and content creation quickly and seamlessly, including ideas for captions; at the same time, they may dynamically explore the context in a way that fits with brand messaging and design constraints (Liu et al., 2025).

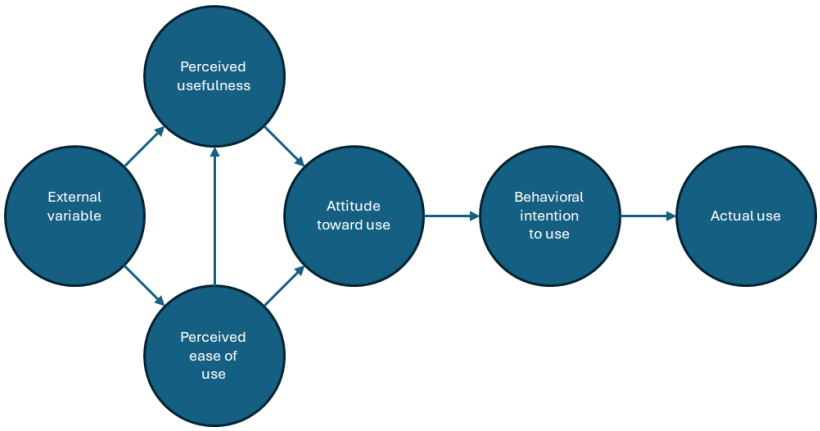
Meanwhile, Canva can be an excellent resource to integrate in EFL classrooms for promoting language skill acquisition, encouraging student involvement, and creating student-centered learning environments. Teachers can produce visual materials, vocabulary games, interactive exercises, and stories that suit students' diverse learning styles and proficiency levels (Emiliyana et al., 2024). In addition, students can create digital comics, visual essays, and presentations which helps strengthen their English communication skills and build confidence (Priyatna, 2023). Canva also promotes a collaborative leaning environment where students can post their co-created projects and receive feedback from fellow students and/or teachers, hence enhancing teamwork and communication abilities (Gehred, 2020). For example, teachers and students join projects via Canva's file-sharing feature to leave comments or assist with grammatical or spelling errors. As a web-based tool, Canva's accessibility and user-friendliness make it an exemplary tool for improving EFL learners' digital literacy and equipping them with essential expertise and creativity to work through the increasingly digitized world.

2.3 Technology Acceptance Model (TAM) and Hypothesis

Grounded in the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980), Technology Acceptance Model (TAM)) aimed to apply the theory of belief-attitude intention-behavior relationship to the context of technology acceptance (Davis,1989). The model has also been extensively utilized as a paradigm for explaining users' technology adoption in educational contexts (Davis et al., 1989; Cheung & Vogel, 2013; Dhingra & Mudgal, 2019; Dahri et al., 2023). The model tries to identify ''the determinants of computer acceptance that are general, capable of explaining user behavior across a broad range of end-user computing technologies and user populations, while at the same time trying to be parsimonious and theoretically justified'' (Davis et al., p. 985). Therefore, the claimed determinants in TAM include individual's attitude (ATT), behavioral intention (BI), and actual use (AU) of a specific technology which are affected by two key factors known as perceived usefulness (PU) and perceived ease of use (PEOU).

Figure 1

TAM model (Davis, 1989)



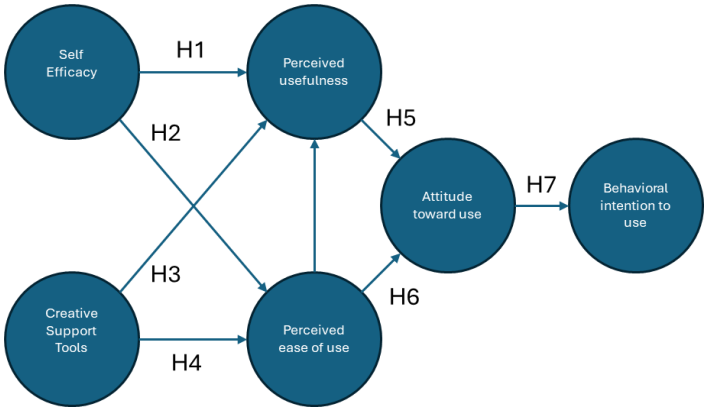
Even though TAM is widely used, it has not been able to fully explain all the things that affect users' acceptance of technological systems because it does not clearly include antecedent or external variables that influence PEOU and PU (Dishaw & Strong, 1999). Davis (1989) also suggested that studies should investigate other factors that could affect PEOU, PU, and AU to enhance its explanatory power and applicability to a wider array of technology implementations in different contexts. Overtime, TAM has undergone several extensions and modifications by including external variables, such as social influence processes (i.e. perceived subjective norms, perceived organizational support), cognitive instrumental processes (i.e. output quality, job relevance, task-technology fit), system characteristics or system design features (i.e. perceived system quality, functionality), individual traits (personal innovativeness, self-efficacy, confidence), and organizational context (i.e. facilitating conditions, broader culture). For example, Dishaw and Strong (1999) proposed an extension of the TAM by integrating task characteristics and technology characteristics. The findings indicated that the new model provides more explanatory power than the original one in clarifying the reasons behind users' adoption of information technologies in task-oriented contexts.

Given the flexibility of the Technology Acceptance Model (TAM) (Davis, 1989) and the variability of specific external factors which can be incorporated within the TAM in different research contexts as well as different study objectives, the author would like to explore two external factors, including self-efficacy (SE) derived from Bandura's (1977) Social Cognitive Theory and creativity support tools (CSTs) based on Shneiderman's (2000) Creativity Support Tool framework. These factors are suitable considering the project's focus on self-guided use of AI tools and are anticipated to have significant impact on students' attitudes and behavioral intention (BI) in incorporating AI tools such as ChatGPT and Canva into their projects. Therefore, a new model was proposed as illustrated in figure 2, and the following factors are investigated in total:

(1) Self-efficacy (SE)
(2) Creativity support tools (CSTs)
(3) Perceived usefulness (PU)
(4) Perceived ease of use (PEOU)
(5) Attitude (Att) toward the AI use
(6) Behavioral intention (BI) for Actual use (AU)

Figure 2

Proposed Theoretical Framework



2.4 Self-efficacy (SE)

Self-efficacy (SE), according to Bandura (1977), is characterized as an individual's belief that one can effectively do the things that will lead to desired outcomes. This belief influences how they act, how hard they try, and how long it will last when coping with difficult situations. In the realm of technology, SE is regarded as significant factor that reflects an individual's self- assessment of their ability to accomplish a specific task by using technology (Compeau & Higgins, 1995; Teo, 2019). Many researchers have undertaken studies across different contexts, including educational institutions and professional settings, to demonstrate the substantial influence of self-efficacy on factors like Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) within the TAM. They found that individuals with high SE are more likely to see AI tools are useful and easy to use, hence increasing their propensity to utilize them (Venkatesh & Bala, 2008; Dahri et al., 2023). Considering the above impacts of SE on learners' AI tool adoption in their project-based tasks, the following hypotheses were set forth:

H1: SE has a positive effect on PU when employing ChatGPT and Canva.

H2: SE has a positive effect on PEOU when employing ChatGPT and Canva.

2.5 Creativity support tools (CSTs)

Unlike self-efficacy, which pertains to an individual's self- assessment of competence in applying AI tools, Shneiderman (2000)'s concept of creative support tools (CSTs) indicates the degree to which AI tools such as ChatGPT and Canva help individual express ideas creatively across diverse domains. Such technologies function not just as tools but as comprehensive platforms that automate routine tasks, improve efficiency, and promote innovative thinking (Secretan, 2011). CSTs do not hinder users' innovation; rather, they highlight the integration of human and machine contributions. Shneiderman (2000) argued that digital tools can scaffold creative work through various features like idea visualization, ideation prompts, reflection, rapid feedback, and design exploration, which might impact users' acceptance and effectiveness of technology. Even though the original model did not explicitly define PEOU or PU, the objectives of his framework as well as acknowledged challenges are intricately connected to the adoption and effective utilization of these tools by individuals. Simultaneously, concurrent research on extending TAM for Web 2.0 collaborative platforms has conceptualized system-level features such as compatibility, resources availability, and peer influence facilitations as external variables that significantly affect PEOU and indirectly PU.

This shapes users' attitudes, behavioral intentions, and actual use in project- based learning environments (Cheung & Vogel, 2013). Hence, the following hypotheses are proposed:

H3: Creativity support tools (CSTs) have a positive effect on Perceived Usefulness (PU).

H4: Creativity support tools (CSTs) have a positive effect on Perceived Ease of Use (PEOU).

2.6 Perceived usefulness (PU)

In the Technology Acceptance Model (TAM), Perceived Usefulness (PU) is considered as a crucial determinant to understand individual's overall attitude, defined as "the degree to which a person believes that using a particular technology would enhance his or her task performance" (Davis, 1989, p.320). It reflects the idea that users with higher willingness to use technology would attain favorable outcomes in accomplishing their goals or tasks such as improving performance, efficiency, and overall productivity. Numerous empirical research demonstrates the positive impacts of PU on users' attitude and behavioral intentions toward the adoption of new technologies (e.g., Davis et al., 1989; Dishaw & Strong, 1999) in educational contexts (e.g., Cheung & Vogel, 2013; Dahri et al., 2023). Therefore, it is probable that PU will result in students' positive perception of ChatGPT and Canva for their class projects.

H5: Perceived usefulness (PU) has a positive effect on students' attitude (Att) toward the use of ChatGPT and Canva.

2.7 Perceived ease of use (PEOU)

PEOU refers to "the degree to which a person believes that using a particular system would be free of effort" (Davis,1989, p.320). It is more likely to focus on the mental aspects, including cognitive resources and mental energy that people use when engaging with technology. In this study, PEOU means how easily Business English students can use and integrate ChatGPT and Canva into their project activities. Like PU, perceived ease of use (PEOU) is another key antecedent that shapes users' attitude towards a new technology. Consequently, PEOU drives their future intentions, which further manifest themselves in actual usage across various fields including education (Davis et al. 1989; Dhingra & Mudgal, 2019). Based on this, the following hypothesis is set forth:

H6: Perceived ease of use (PEOU) has a positive effect on students' attitude (Att) towards ChatGPT and Canva use

2.8 Attitude (Att) towards the AI use

The model includes attitude towards use (Att), which directs users' future behaviors toward AI tools. Attitude reflects the positive-negative evaluation, opinions, and sentiments with respect to performance of any technology, whereas it is theorized to mediate the effects of PEOU and PU on BI (Davis, 1989). Previous research by Wong et al. (2013) applied the TAM model to explore Malaysian student teachers' attitude toward computer implementation in classrooms. The findings supported that students' perceptions of usefulness (PU) and ease of use (PEOU) significantly affected their attitudes toward computer integration, which subsequently influenced their adoption intention towards learning activities. Shahzad et al. (2024) surveyed 320 Chinese university students and extended the TAM by including perceived intelligence to identify factors influencing students' intention to use ChatGPT. They found that perceived intelligence was a significant determinant in the adoption process and established a positive association between students' attitudes and behavioral intentions toward ChatGPT. These findings suggest that when students perceive Chat GPT as highly intelligent and capable of providing quality responses, their likelihood of using the tool increases. Based on these insights and the research objectives, Att towards AI use serves as a critical emotional factor that bridges cognitive appraisal, such as PEOU and PU, and behavioral intention (BI). The following hypothesis is therefore proposed:

H7: Students' attitude (Att) toward ChatGPT and Canva use has a positive effect on students' behavioral intention (BI) for actual use.

2.9 Behavioral intention (BI) for Actual use (AU)

To understand students' attitude towards using AI tools, the entire student adoption process cannot help but encompass covering behavioral intention (BI) for actual use (AU). Furthermore, due to its predictive design, the TAM includes BI as the factor of acceptance to predict actual technological use. Davis et al. (1989) asserts that BI reflects the degree to which an individual intends to use a particular system, serving as a strong determinant of AU behavior. Multiple studies have confirmed that learners' attitudes toward technology use significantly impact their BI and that a positive attitude may lead to greater commitment and use (Venkatesh & Davis, 2000). In the field of EFL, Coşkun and Solmaz (2024)'s study investigated the perceptions and attitudes of 650 high school English learners towards a mobile language learning application. They discovered that learners' intention to continue using mobile tools increases when they regard digital tools as important and advantageous for their various study goals, including for language practice (i.e., practicing, building English language abilities) and for project completion. Based on this foundation, the study investigates whether students' attitudes towards ChatGPT and Canva in their projects influence their BI to lead to actual use (AU) of these tools.

3. Methodology

3.1 Research design

The current study employed a quantitative methodology to investigate students' attitudes and behavioral intention regarding the use of ChatGPT 3.5 and Canva in a project-based learning activity, especially within a Selling Skills course at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). The online survey method was chosen to collect data due to its cost-effectiveness and time efficiency in regional distribution (Kurfali et al., 2017).

3.2 Sampling and data collection

The subjects of this study were third year university students majoring in Business English at HUFLIT. All participants were enrolled in the author's 12-week Selling Skills course, which was a compulsory EFL component of the university curriculum. Their English proficiency levels ranged from intermediate (B1) to upper-intermediate (B2) according to the Common European Framework of Reference for Languages (CEFR). The survey questionnaires were prepared in English and then distributed electronically via Google from which could be accessible from participants' personal devices. Initially, 150 students from three classes taught by the author were invited to participate in the survey at the end of the course in June 2025; however, only 136 of those students eventually answered the survey. Even though the results may have limited generalizability, it is convenient for a specific group of respondents and relevant to the items that are being measured (Shahzad et al., 2022). In terms of demographic information, there were 102 females (75 %), 31 males (22.79%), and 3 individuals (2.21%) who declined to disclose their gender. All participants are aged between 20 and 22 years and voluntarily consented to participate in the survey in compliance with institutional ethical standards.

3.3 Research Instrument

To investigate constructs associated with individuals' attitudes and behavioral intentions toward AI, particularly focusing on ChatGPT and Canva, participants were asked to provide responses to two essential demographic questions—gender and age—alongside twenty comprehensive scale items utilizing a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree): Self-efficacy (3 items) adapted from Murphy et al. (1989), creativity

support (4 items) adapted from Cherry, E., & Latulipe, C. (2014), perceived usefulness (3 items) adapted from Davis (1989) and Venkatesh et al. (2003), perceived ease of use (3 items) adapted from Davis (1989) and Venkatesh & Davis (2000), attitude (4 items) toward the AI use adapted from Davis (1989), Venkatesh et al., (2003), and Venkatesh & Davis (2000), and behavioral intention (3 items) adapted from Venkatesh et al., (2003), and Venkatesh et al., (2012). These items were derived from previously well researched studies which indicated a high degree of reliability. Table 1 below shows variables, statements, and sources of the items.

Table 1

Items used to measure TAM constructs

Factors	Items	Statements	Reference
Self-efficacy (SE)	SE1	I was confident in my ability to use different features of ChatGPT and Canva to complete my project tasks.	Murphy et al. (1989)
	SE2	I felt capable of solving problems I met while using ChatGPT and Canva during the project.	
	SE3	I was confident in learning how to use new features of ChatGPT and Canva on my own.	
Creativity support tools (CSTs)	CST1	The combination of ChatGPT and Canva in content generation and visual design helped to enhance my creativity during the project.	Cherry, E., & Latulipe, C. (2014)
	CST2	ChatGPT and Canva allowed me to come up with creative ideas for my project.	
	CST3	ChatGPT and Canva supported my team's collaboration, which generated highly creativity.	
	CST4	ChatGPT's and Canva's diverse features helped me produce creative designs and posts.	
Perceived Usefulness (PU)	PU1	Using ChatGPT and Canva saved my time on project tasks.	Davis (1989), and Venkatesh et al. (2003)
	PU2	ChatGPT and Canva helped to enhance the overall efficiency of my project.	
	PU3	ChatGPT and Canva helped me achieve better results in PBL activities.	
Perceived Ease of Use (PEOU)	PEOU1	I found it easy to navigate the interfaces of ChatGPT and Canva.	Davis (1989), and Venkatesh & Davis (2000)
	PEOU2	I could learn how to use ChatGPT and Canva for my project quicky.	
	PEOU3	It was clear and understandable to interact with ChatGPT and Canva.	
	Att1	I enjoyed integrating ChatGPT and Canva into my project.	Davis (1989), Venkatesh et al.,

Attitude to ChatGPT and Canva use (Att)	Att2	I believe using ChatGPT and Canva was a good idea for my project.	(2003), and Venkatesh & Davis (2000)
	Att3	ChatGPT and Canva are great for projects that involve collaboration.	
	Att4	I have a positive attitude toward ChatGPT and Canva use for similar projects.	
Behavioral Intention to use ChatGPT and Canva (BI)	BI1	I intend to use ChatGPT and Canva for any PBL assignments in the future.	Venkatesh et al., (2003), and Venkatesh et al., (2012)
	BI2	I would recommend ChatGPT and Canva to peers for project collaboration.	
	BI3	I plan to incorporate ChatGPT and Canva into my regular study routines	

4. Results and data analysis

After the data had been gathered online through Google forms and downloaded in a Microsoft Excel file in the .csv format, it was then integrated into the SmartPLS software version 3.0 for the purpose of statistical analysis. Partial Least Square-Structural Equation Modelling (PLS-SEM) was employed to investigate the conceptual model in three different areas: first, the measurement model, second, the structural mode, and third, hypothesis testing. According to Hair et al. (2017), PLS-SEM is a suitable technique for the investigation of complicated cause-effect structures, hence enhancing the comprehension of inter-relationships between observable and latent variables. It might also be utilized in research pertaining to acceptance and adoption. Therefore, the requisite steps must be undertaken to evaluate the reliability and validity of the construct, beginning with an assessment of the construct's internal consistency or reliability (outer loading coefficient, Cronbach's Alpha, CR), followed by an examination of the construct's convergent validity (AVE coefficient) and discriminant validity (Fornell-Larcker criteria). Subsequently, the linear structural model SEM should be analyzed using the variance inflation factor (VIF), R square value (R^2), path coefficient (β), communality, Cohen's f^2, Q^2 and SRMR. Finally, P value was employed to validate the proposed hypotheses (Hair et al., 2017).

4.1 Measurement model

The evaluation of the measurement model was validated through composite coefficients which are shown in Table 2 and are as follows:

Table 2
Summary of coefficients in the PLS-SEM model

Factors	Items	Factor Loading	Cronbach's Alpha	Composite Reliability (CR)	Average Variance Extracted (AVE)	VIF
Self-efficacy	SE1	0.919	0.914	0.946	0.854	2.977
	SE2	0.925				3.275
	SE3	0.927				3.289

Creativity support tools	CST1	0.906	0.855	0.903	0.699	3.168
	CST2	0.776				1.692
	CST3	0.822				1.913
	CST4	0.836				2.274
Perceived Usefulness	PU1	0.880	0.855	0.912	0.775	2.129
	PU2	0.862				1.989
	PU3	0.898				2.310
Perceived Ease of Use	PEOU1	0.899	0.861	0.915	0.783	2.432
	PEOU2	0.845				1.903
	PEOU3	0.909				2.500
Attitude to ChatGPT and Canva use	Att1	0.844	0.888	0.923	0.750	2.576
	Att2	0.922				3.794
	Att3	0.899				3.216
	Att4	0.794				1.717
Behavioral Intention to use ChatGPT and Canva	BI1	0.857	0.784	0.869	0.689	1.412
	BI2	0.822				1.873
	BI3	0.810				1.907

* Source: SmartPLS output

The reliability or consistency of a scale or questionnaire was initially evaluated using Cronbach's Alpha (CA). Hair et al. (2017) assert that a Cronbach's alpha (CA) value of 0.70 or above signifies good reliability, whereas values over 0.80 reflect qualified reliability. In table 2, all measured constructs return Cronbach's Alpha values of being greater than 0.70. Especially, the coefficients are as follows: Self-efficacy (0.914), Creativity Support Tools (0.855), Perceived Usefulness (0.855), Perceived Ease of Use (0.861), Attitude (0.888), and Behavioral Intention (0.784). Furthermore, the majority of item loadings on their associated factors are above 0.7, with the lowest being 0.776, representing strong relationships between items and their respective factors. These results confirm that the items within each construct are reliable and consistently measure the intended latent variables.

Likewise, the composite reliability (CR) scores were also examined since they provide a more precise measure of internal consistency than Cronbach's alpha by accounting for the differing reliabilities of distinct indicators through weighted loadings (Hair et al., 2017). The convergent reliability (CR) value for each construct was greater than 0.8, with values ranging from 0.869 (Behavioral Intention) to 0.946 (Self-efficacy). Accordingly, the values bring robustness and reliability to the measurement model, which supports the consistency of the constructs.

The constructs also showed notable convergent validity, as denoted through the Average Variance Extracted (AVE). This value must surpass 0.50 to demonstrate that each construct accounts for over half of the variance of their respective indicators (Hair et al., 2017). In this study, all of the component AVE values were greater than the required cut-off, varying between 0.689 (Behavioral Intention) and 0.854 (Self-efficacy), thereby confirming that the constructs sufficiently capture the variance of their respective indicators, thus affirming convergent validity within the measurement model.

Table 3 shows the discriminant validity of the constructs based on the Fornell–Larcker criterion. Hair et al. (2017) emphasized the importance of highlighting the diagonal values which reflects the square root of the Average Variance Extracted (AVE) for each construct. Especially, the square root of the AVE on the diagonal values should be higher than the off-diagonal values for each pair. In this case, Attitude (0.866), Behavioral Intention (0.830), Creativity Support Tools (0.836), Perceived Ease of Use (0.885), Perceived Usefulness (0.888), and Self-efficacy (0.924) consistently surpass their respective inter-construct correlations in the rows and columns. This evidence confirms that the discriminant validity of the constructs is achieved, ensuring the factors are empirically distinct from one another.

Table 3

Discriminant validity

	Att	BI	CSTs	PEOU	PU	SE
Att	**0.866**					
BI	0.779	**0.830**				
CSTs	0.800	0.691	**0.836**			
PEOU	0.875	0.710	0.773	**0.885**		
PU	0.790	0.690	0.748	0.781	**0.880**	
SE	0.882	0.684	0.755	0.847	0.738	**0.924**

* Source: SmartPLS output

4.2 Structural Model

In PLS-SEM, the structural model represents proposed relationships between different variables (Hair et al., 2017). To detect potential multicollinearity across the constructs, the Variance Inflation Factor (VIF) is examined and should be between 0.20 and 5.00 (Hair et al. (2017), where values below 5 indicate that multicollinearity does not pose a serious problem. According to the measurement model results shown in table, all items have VIF values ranging between 1.412 and 3.794, which are within acceptable range.

Table 4 presents the analysis of the coefficient of determination (R^2), shedding light to the model's explanatory power. Adjusted R^2 values were taken into consideration since they measure the number of predictors of the model while providing a more cautious assessment of explanatory accuracy. In this study, the R2 adjusted value was 0.791 for Attitude and 0.755 for Perceived Ease of Use, both indicating strong explanatory power. Meanwhile, the adjusted R^2 value for Perceived Usefulness (0.623) and Behavioral Intention Adjusted (0.604) fall within the moderate range. Overall, the results confirm the strength and substantial predictive power of the aforementioned model, with Attitude (Att) and Perceived Ease of Use (PEOU) being well explained by their predictors, while Behavioral Intention (BI) and Perceived Usefulness (PU) are moderately explained.

Table 4

R and R^2 Validation

	R Square	R Square Adjusted	Results
Att	0.794	0.791	Strong
BI	0.607	0.604	Moderate
PEOU	0.759	0.755	Strong
PU	0.629	0.623	Moderate

* Source: SmartPLS output

For overall fit of the structural model (PLS–SEM), the Communality index was assessed. Hair et al. (2017) say that the Communality index is the same as AVE in the PLS model and that it should be more than 0.5 for the model to be declared fit. Table 2 shows that the structural model has AVE values higher than 0.5 for all of the factors.

The study further investigated Cohen (2013)'s effect size index (f2) to classify their influence into 3 groups: (1) small effect size with f^2 less than 0.1, (2) medium with f2 between 0.25 and 0.40, and (3) large effect size with f^2 greater than 0.4 . According to PLS results shown in table 7, practical significance of the relationship between variables in the model fell in two groups: H1, H3, H4, and H5 exhibit a small effect, while H2, H6, and H7 have a substantial effect with f^2 values over 0.4. In addition, Table 5 contains the values of the Q^2 coefficient, which represents the level of prediction accuracy of the structural model. In other words, these values help to explain how well the model predicts the observed values. According to the findings, Att (0.582) and PEOU (0.583) demonstrate high predictive relevance while PU (0.477) and BI (0.378) show medium relevance (Hair et al. 2017).

Table 5

Q^2 Coefficient Analysis Results

	Saturated Model	Estimated Model
SRMR	0.081	0.090
d_ULS	1.367	1.697
d_G	1.138	1.248
Chi-Square	701.037	740.065
NFI	0.741	0.726

* Source: SmartPLS output

An additional critical metric for evaluating the overall fit of the structural model in SEM is Standardized Root Mean Square Residua (SRMR). This measure is used to assess how well the observed data matches the predicted values of the model. In accordance with Hu and Bentler (1998)'s guide and the SRMR results in table 6, the proposed structural model in this study fell within the tentative acceptance range (SRMR <0.10). In general, this study has conclusively shown that the PLS model effectively aligns with the empirical data based on the outcomes analyzed above.

Table 6

Model Fit Summary

	SSO	SSE	Q^2 (=1-SSE/SSO)
Att	544.000	227.605	0.582
BI	408.000	253.676	0.378
CSTs	544.000	544.000	
PEOU	408.000	169.948	0.583
PU	408.000	213.523	0.477
SE	408.000	408.000	

* Source: SmartPLS output

4.3 Hypotheses testing

To test the strength and direction of relationships between latent variables, path coefficients or standardized beta coefficients (β) were assessed, yielding values between 0 and 1 (Hair et al., 2017). The higher path coefficient (β) means a stronger relationship, while values closer to zero mean weaker effects. Through the results shown in table 7 and in figure 3, the strongest path was between Attitude and Behavioral Intention (β = 0.779), showing that students who have positive attitude towards the use of ChatGPT and Canva in their projects are more likely to intend to use such tools in the future. Among the factors influencing Attitude, PEOU showed a more significant impact with β of 0.660 in contrast to PU, which exhibited a β of 0.275. In relation to external factors, CSTs was found to have clear influences on both PU (β = 0.444) and PEOU (β = 0.310), while SE showed a more significant influence on PEOU (β = 0.612) than on PU (β = 0.403). This evidence suggests that PU and PEOU could mediate the effects of CSTs and SE on students' Att and BI regarding the use of ChatGPT and Canva in PBL.

The final phase was validating the hypotheses posited in this study by performing the bootstrapping method, with a repeated sample size of 5000 observations. Hypotheses were accepted only if the significance criterion of p-value < 0.05 was achieved (Hair et al., 2017). As shown in table 7, all seven tested hypotheses demonstrated significant relationships, with p-values substantially below 0.05; thereby, suggesting strong statistical support for the proposed paths.

Table 7

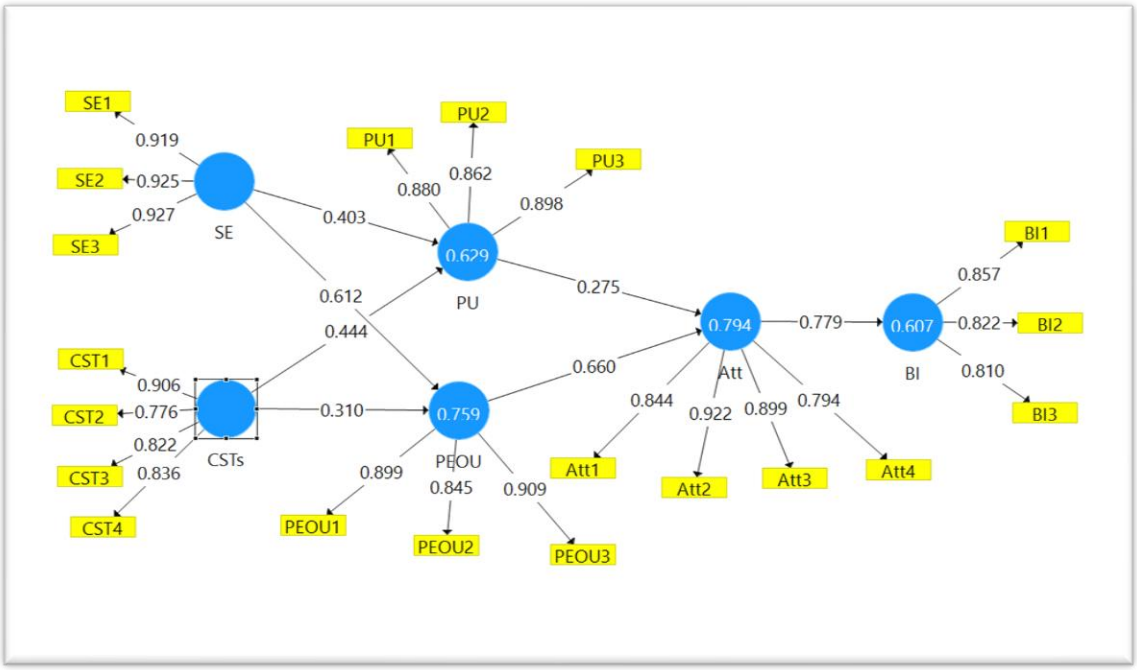
Results of Relationship Testing between Variables

Hypothesis	Relationship	Path coefficient	P-Value	f^2	Result
H1	SE→PU	0.403	0.000	0.188	**Accepted**
H2	SE→PEOU	0.612	0.000	0.668	**Accepted**
H3	CSTs→PU	0.444	0.000	0.228	**Accepted**
H4	CSTs→PEOU	0.310	0.000	0.172	**Accepted**
H5	PU→Att	0.275	0.004	0.144	**Accepted**
H6	PEOU→Att	0.660	0.000	0.826	**Accepted**
H7	Att→BI	0.779	0.000	1.545	**Accepted**

* Source: SmartPLS output

Figure 3

Estimated Model Results using SMART PLS-SEM



5. Discussion

This study utilized the Technology Acceptance Model (TAM) framework, incorporating two additional external factors (Students' Self-Efficacy and AI's Creative Support). The aim was to explore students' attitudes and their behavioral intentions after using ChatGPT and Canva to fulfil their project-based learning activities which were related to generating content and designing product promotion posts within a Marketing and Advertising course. Therefore, a conceptual model was proposed and statistically verified based on data collected from 136 Business English students at English Department of HUFLIT. The author used an online survey-based questionnaire technique to collect feedback on using the ChatGPT and Canva and distributed it to the intended audience. Followings are the remarkable results found from the investigation:

First, students' self-efficacy was found to have significant influence on both perceived usefulness ($\beta = 0.403$, $p = 0.000$) and perceived ease of use (PEOU) ($\beta = 0.612$, $p = 0.000$) of AI technologies in project-based learning (PBL) activities. This finding aligns with previous study of Venkatesh and Bala (2008), highlighting that students who perceive themselves as capable and confident in using digital tools are more likely to consider ChatGPT and Canva both useful and easy to use. In addition, when students believe they can use AI and design tools, they are more inclined to regard AI beneficial for enhancing creativity, productivity, and learning outcomes, so substantially shape students' attitudes and future intentions (Dahri et al.,2023). In other words, PU and PEOU help to mediate the relationship between students' self-efficacy and their attitude.

Similarly, CSTs demonstrated as a significant determinant of PU and PEOU, displaying $\beta = 0.444$, $p = 0.000$ and $\beta = 0.310$, $p = 0.000$, respectively. Within the scope of the study, Business English students were tasked with creating product promotion posts on social media platforms like Facebook or Instagram, AI tools such as ChatGPT and Canva provided a variety of features and resources to enhance their creativity. For example, ChatGPT may help with idea generation and brainstorming slogan or contents for social media postings. Meanwhile, Canva offers pre-designed templates, images, layouts, and other components for the creation of visuals. These platforms also include collaborative features through which students can

flexibly work, share, co-create, give feedback, and revise collectively (Andriyanto et al., 2021; Gehred, 2020). This somehow reflects the growing importance of such creativity support tools in educational settings, particularly for project-based learning, where innovation and creative problem-solving are highly valued. Hence, this evidence is in line with Shneiderman (2000) and Cheung and Vogel (2013)' s research, indicating that tools designed to foster creativity enhance users' perceived utility and usability.

In the extended TAM proposed by the author, perceived usefulness (PU) and perceived ease of use (PEOU) were dependent variables, influenced by two external factors: self-efficacy (SE) and creativity support (CS). In TAM, both PU and PEOU were essential indicators of predicting students' attitudes towards using ChatGPT and Canva. To be specific, PU had a significant positive effect on students' attitudes ($\beta = 0.275$, $p = 0.004$), indicating that when students perceive the usefulness of using AI tools such as ChatGPT and Canva for tasks like project and assignment completion, their favorable attitudes towards these tools are subsequently amplified. This outcome corresponds with previous studies on the implementation of the original TAM model (Davis, 1989; Cheung & Vogel, 2013). Simultaneously, PEOU exhibited a more notable impact on attitudes ($\beta = 0.660$, $p = 0.000$), suggesting that students prioritize the simplicity and usability of the tools when forming their overall impression. This result is also consistent with results of Davis et al. (1989) and Dishaw and Strong (1999) who highlighted the fact that PEOU is more critical than PU for AI tools integration. Generally, even though students value usefulness, the ease of using ChatGPT and Canva has a greater impact on forming positive attitudes. This phenomenon can be due to the relatively new integration of AI tools in education, where user-friendliness, such as Canva's drag-and-drop functionalities and ChatGPT's natural language interactions accelerates acceptance.

Finally, the results displayed a significant correlation between attitude and behavioral intention towards the use of ChatGPT and Canva ($\beta = 0.779$, $p = 0.000$), which is also the most substantial relationship within the model. The term "attitude" generally reflects students' self-evaluation and emotional disposition when engaging with AI tools. In the context of this study, students clearly showed favorable attitudes towards the use of ChatGPT and Canva in their project. Attitude is also regarded as the key factor in TAM which mediates the effect of PU and PEOU on students' behavioral intention to utilize AI technology. When students cultivate positive attitudes toward ChatGPT and Canva, there is a likelihood that they will continue to integrate these tools into their academic tasks and other future project-based learning activities. This confirms the findings of Wong et al. (2013), who found that positive attitudes directly drive the intention to use technology in higher education settings. Therefore, it is important to foster positive initial impressions and satisfaction among students as it will subsequently lead to sustained AI adoption behavior.

6. Conclusion and Implication

The purpose of this study was to explore university Business English students' attitudes and their behavioral intentions regarding the use of ChatGPT and Canva in a marketing project, as part of the coursework assessment. The author utilized a quantitative approach along with enhanced TAM as theorical framework to gather valuable insights. All variables and hypotheses have been examined and confirmed by PLS-SEM analysis, shedding light on their relevance and influence in the context of the targeted sample. A significant result found was students' positive attitudes towards the implication of AI tools for specific tasks, which in turn impact their behavioral intentions (Davis, 1989). The study also found that students value the functional benefits as well as the simplicity of integrating AI. Especially for social media product promotion projects, the combination of ChatGPT and Canva in content generation and

visual design features enabled students to ideate, design, and create professional-quality marketing content, thereby fostering engagement, skill development, creativity, and productivity. Furthermore, examining the two external factors, namely self-efficacy (SE) and creativity support (CSTs) contributed to better understanding the fundamental constructs of TAM and how they play a role in shaping students' attitudes and behavioral intentions.

The current study presents some practical insights for various stakeholders in educational institutions who seek to integrate AI and design tools into PBL activities. Given that students express positive attitudes towards AI tools, educators, institutions, and developers should nurture that advantage by providing sufficient support for effective AI implementation; at the same time, leveraging their potential impact on language education. This emphasizes the need of providing training sessions, guided tutorials, workshops, and structured project instructions to familiarize students with AI tools and build their confidence, which can enhance their perceived usefulness as well as perceived ease of use (Venkatesh & Bala, 2008; Dahri et al., 2023). Institutions should consider integrating AI tools into curricula by establishing explicit instructions for educators. This will enhance educators' confidence and motivation to apply AI tools in their classes, hence increasing student engagement with study and real-world activities. Additionally, the effectiveness of AI tools in the classroom can be further enriched if educators invest in designing learning environments that are authentic, collaborative, and creative. Educators should consider numerous factors when formulating pedagogically effective methodologies. This includes not only technical and ethical dimensions of these tools but also pedagogical aspects such as learning objectives, instructional strategies, and their roles as facilitators to optimize the usefulness of AI. These proposed implications align with the contemporary educational landscape, emphasizing the importance of technology, particularly AI tools, in facilitating language acquisition across various learning settings.

7. Limitation and Future Research

This study, like any research, was not without its limitations. Firstly, the data was collected from a specific student population, thus constraining the generalizability of the findings. This is due to the nature of the investigation's concentration solely on Business English students for project assignments within the author's Advertising and Marketing course. Researchers are advised to further investigate students' attitudes and behavioral intentions by expanding larger, and more randomized samples from various classes, faculties and institutions. This can also be implemented in other educational disciplines, including course subjects, PBL tasks, or assignments through collaboration with other instructors. Secondly, the study mainly focused on two typical tools – ChatGPT and Canva- which may not reflect the comprehensive landscape of emerging AI tools, especially in educational settings. Therefore, there should be more research looking into the integration of alternative technologies as well as their impact on user acceptance. Third, the present study made use of quantitative approach, utilizing self-reported measures which may introduce response bias or inaccuracies. Future research should incorporate mixed techniques with interviews or focus groups to achieve more objective assessments as well as to gain a more profound knowledge of users' opinions, their potential implementation challenges, and their expectations. A comparative study is recommended to examine the distinct implementations of AI-driven and traditional project-based learning tools in EFL classes. Lastly, regarding the proposed model incorporating two external variables—self-efficacy and creative support—research findings revealed that most of variables had direct and indirect correlations with Attitude and Behavioral Intention. Therefore, future studies should incorporate additional variables related to project-based learning activities, such as AI experiences and collaboration skills, while recontextualizing them within the Technology Acceptance Model to clarify their roles more distinctly.

References

Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. In J. Kuhl & J. Beckmann (Eds.), Action control: From cognition to behavior (pp. 11–39). Springer-Verlag.

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Prentice-Hall.

Andriyanto, O. D., Hardika, M., & Sukarman, S. (2021). Canva as a media for exploration of Indonesian tourism knowledge in BIPA learning. In Proceedings of the International Joint Conference on Arts and Humanities (IJCAH 2021) (Vol. 633, pp. 600–605). https://doi.org/10.2991/assehr.k.211223.034

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Belda-Medina, J. (2021). ICTs and project-based learning (PBL) in EFL: Pre-service teachers' attitudes and digital skills. International Journal of Applied Linguistics & English Literature, 10(1), 1–8. https://doi.org/10.7575/aiac.ijalel.v.10n.1p.1

Belda-Medina, J., & Goddard, M. B. (2024). AI-driven digital storytelling: A strategy for creating English as a foreign language (EFL) materials. International Journal of Linguistics Studies, 4(1), 40–49. https://doi.org/10.32996/ijls.2024.4.1.4

Browning-Samoni, L. (2024). AI and fashion: Student perspectives on the application and ethical use of various forms of generative artificial intelligence (GAI) in a fashion context (Doctoral dissertation proposal, Iowa State University). Iowa State University.

Campbell, C. (2014). Problem-based learning and project-based learning. *Australian Council for Educational Research–ACER*.

Cherry, E., & Latulipe, C. (2014). Quantifying the creativity support of digital tools through the Creativity Support Index. ACM Transactions on Computer-Human Interaction (TOCHI), 21(4), Article 21, 1–25. DOI: https://doi.org/10.1145/2617588

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers & education*, *63*, 160-175.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. MIS Quarterly, 19(2), 189–211.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.

Dahri, N. A., Al-Rahmi, W. M., Almogren, A. S., Yahaya, N., Vighio, M. S., Al-Maatuok, Q., Al-Rahmi, A. M., & Al-Adwan, A. S. (2023). Acceptance of mobile learning technology by teachers: Influencing mobile self-efficacy and 21st-century skills-based training. Sustainability, 15(8), 8514. https://doi.org/10.3390/su15118514

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35(8), 982–1003.

Dhingra, M., & Mudgal, R. K. (2019). Applications of perceived usefulness and perceived ease of use: A review. In 2019 8th International Conference on System Modeling and Advancement in Research Trends (SMART) (pp. 293–298). https://doi.org/10.1109/SMART46866.2019.9117404

Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task-technology fit constructs. Information & Management, 36(1), 9–21.

Simpson, D. J., & Stack Jr., S. F. (Eds.). (2010). *Teachers, leaders, and schools: Essays by John Dewey*. SIU Press.

Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. Journal of University Teaching & Learning Practice, 20(5). https://doi.org/10.53761/1.20.5.02

Emiliyana, M. P., Fitriah, S., & Kurniati, U. (2024). Teacher and student's pedagogical tasks using Canva as media for English language teaching (ELT) in developing creativity for high school. International Journal of English Language Studies, 6(2), 130–136.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.2307/3151312

Gehred, A. P. (2020). Canva. Journal of the Medical Library Association, 108(2), 338–340. https://doi.org/10.5195/jmla.2020.940

Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. Education Sciences, 13(12), 1216. https://doi.org/10.3390/educsci13121216

Monika, M., & Suganthan, C. (2024). A study on analyzing the role of ChatGPT in English acquisition among ESL learners during English language classroom. *Bodhi International Journal of Research in Humanities, Arts and Science*, *8*(2), 75-84.

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2017). A Primer on PartialLeast Squares Structural Equation Modeling (PLS-SEM). 2nd Edn. Thousand Oaks, CA:Sage Publications

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? Educational Psychology Review, 16(3), 235–266.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Kurfali, M., Arifoğlu, A., Tokdemir, G., & Paçin, Y. (2017). Adoption of e-government services in Turkey. Computers in Human Behavior, 66, 168–178. https://doi.org/10.1016/j.chb.2016.09.041

Liang, X., & Luo, J. (2024). Integration of ChatGPT into project-based learning: A course design framework. International Journal of Chinese Language Teaching, 5(1), 29–46. https://doi.org/10.46451/ijclt.20240104

Liu, X., Sun, A., An, P., Ma, T., & Zhao, J. (2025). Influencer: Empowering everyday users in creating promotional posts via AI-infused exploration and customization. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. ACM. https://doi.org/10.1145/3706598.3713309

Miller, M. (2023). *AI for educators: Learning strategies, teacher efficiencies, and a vision for an artificial intelligence future*. Ditch That Textbook.

Ministry of Education and Training. (2024, August 23). *Decision No. 2236/QĐ-BGDĐT on promulgating the plan for key tasks and solutions for the 2024–2025 academic year of the education sector*. https://thuvienphapluat.vn/van-ban/Giao-duc/Quyet-dinh-2236-QD-BGDDT-2024-nhiem-vu-trong-tam-nam-hoc-2024-2025-nganh-Giao-duc-622626.aspx

Murphy, C. A., Coover, D., & Owen, S. V. (1989). Development and validation of the computer self-efficacy scale. *Educational and Psychological measurement*, *49*(4), 893-899.

O'Dea, X., & O'Dea, M. (2023). Is artificial intelligence really the next big thing in learning and teaching in higher education? A conceptual paper. Journal of University Teaching & Learning Practice, 20(5). https://doi.org/10.53761/1.20.5.05

O'Regan, G. (2016). History of artificial intelligence. In Introduction to the history of computing. Springer. https://doi.org/10.1007/978-3-319-33138-6_19

Shahzad, F., Shahzad, M. F., Dilanchiev, A., & Irfan, M. (2022). Modeling the influence of paternalistic leadership and personality characteristics on alienation and organizational culture in the aviation industry of Pakistan: The mediating role of cohesiveness. Sustainability, 14(22). https://doi.org/10.3390/su142215473

Shahzad, M. F., Xu, S., & Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: Mediating roles of perceived ease of use, usefulness, and intelligence. Educational Technology Research and Development, 72(1), 45–64. https://doi.org/10.1186/s41239-024-00478-x

Shneiderman, B. (2000). Creating creativity: User interfaces for supporting innovation. ACM Transactions on Computer-Human Interaction, 7(1), 114–138. https://doi.org/10.1145/344949.344972

Secretan, J. (2011, November). Semi-automated creativity: Software as a creative collaborator. In Proceedings of the 8th ACM Conference on Creativity and Cognition (pp. 447–448). Association for Computing Machinery. https://doi.org/10.1145/2069618.2069743

Stoller, F. (2006). Establishing a theoretical foundation for project-based learning in second and foreign language contexts. In G. H. Beckett, & P. C. Miller (Eds.), *Project-based second and foreign language education: Past, present, and future* (pp. 19-40). Information Age Publishing.

Tang, X., Ding, X., Ma, X., Zhang, S., & Diao, J. (2024). An exploration of project-based learning supported by artificial intelligence. In Proceedings of the 2024 5th International Conference on Big Data and Informatization Education (ICBDIE 2024) (pp. 220–230). Atlantis Press. https://doi.org/10.2991/978-94-6463-417-4_20

Teo, T. (2019). Students and teachers' intention to use technology: Assessing their measurement equivalence and structural invariance. Journal of Educational Computing Research, 57(1), 201–225.

Thomas, J. W. (2000). *A review of research on project-based learning*. https://tecfa.unige.ch/proj/eteach-net/Thomas_researchreview_PBL.pdf

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. Management Science, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178. https://doi.org/10.2307/41410412

Wong, K.-T., Osman, R., Goh, P. S. C., & Rahmat, M. K. (2013). Understanding student teachers' behavioural intention to use technology: Technology acceptance model (TAM) validation and testing. *International Journal of Instruction, 6*(1), 89–104.

Xiao, Y., & Zhi, Y. (2023). An Exploratory Study of EFL Learners' Use of ChatGPT for Language Learning Tasks: Experience and Perceptions. Languages, 8(3), 212. https://doi.org/10.3390/languages8030212

Yufeia, L., Saleh, S., Jiahui, C., & Syed, M. S. S. (2020). Review of the application of artificial intelligence in education. International Journal of Innovation, Creativity and Change, 12, 548–562.

Coşkun, Z., & Solmaz, O. (2024). Exploring High School Students' Acceptance and Use of Mobile Applications for English Language Learning. *Journal of Language Education and Research*, *10*(2), 636-661. https://doi.org/ 10.31464/jlere.1515400

Bionote

Ms. Le Bao Vi is a lecturer and researcher specializing in English language education. With more than 15 years of experience in teaching English and English for Specific Purposes, she is interested in language teaching methodology, communication skills, and professional English training. She has participated in various academic and professional development activities and is committed to enhancing students' language competence and career readiness.

#40 INVESTIGATING ENGLISH-MAJOR UNDERGRADUATES' READINESS FOR AI-SUPPORTED ACADEMIC WRITING: A MULTIDIMENSIONAL PERSPECTIVE

Tran Phuong Anh [1], & Ngo Thi Thu Hien[2]

Abstract

The integration of Artificial Intelligence (AI) tools in academic writing has expanded rapidly, bringing both opportunities and challenges to the field of teaching English as a foreign language (EFL). Although the use of AI has become increasingly common among Vietnamese university students, empirical research on how English majors are prepared to engage with AI in academic writing remains limited. To address this gap, the present study adopts a mixed-methods approach to investigate English-major undergraduates' readiness and their intention to use AI-supported tools for academic writing, drawing on a multidimensional construct comprising (1) cognitive-metacognitive readiness, (2) linguistic readiness, (3) ethical and responsible awareness, and (4) technology readiness. A questionnaire survey was administered to 93 third-year English majors at Ho Chi Minh City University of Foreign Languages – Information Technology. Quantitative data were analyzed using descriptive statistics, correlation, and multiple regression. The results indicate that not all readiness dimensions significantly predict students' intention to use AI tools. In particular, technology enablers, a sub-dimension of technology readiness, emerged as the only significant predictor, while other dimensions played a less decisive role. In addition, semi-structured interviews were conducted to explore students' perceptions of AI-supported writing. Overall, the findings suggest that positive technological beliefs and supportive conditions outweigh perceived barriers in shaping intentions to use AI. The study contributes to TESOL research on AI-enhanced language learning and offers pedagogical implications for the responsible integration of AI tools in English academic writing instruction at Vietnamese universities.

Keywords: AI-supported academic writing; English-major undergraduates; multi-dimensional readiness; technology readiness; TESOL

1. Introduction

In recent years, artificial intelligence (AI) has improved many aspects of academic writing, including idea creation, content organization, literature synthesis, data management, editing, and ethical compliance (Khalifa & Albadawy, 2024). For English undergraduates who are required to learn and practise advanced writing skills, AI tools support them not only to produce academic texts but also to reduce human errors (Setiana & Chamalah, 2025).

According to Akanda and Talukder (2026), AI tools play a significant role in academic writing by providing immediate feedback on grammar, spelling and structure, which may help learners with their writing. However, many students are increasingly dependent on AI usage, which leads to a decline in their critical and independent thinking and writing abilities (Liu et al., 2023). In addition, there are challenges to academic integrity, such as plagiarism (De

[1] *Master of English Language Studies, Independent Researcher, Ho Chi Minh City, Vietnam; Email:phuonganh.tr1089@gmail.com*
[2] *Master of English Language Studies, Independent Researcher, Ho Chi Minh City, Vietnam; Email:friendship200561@gmail.com*

Angelis et al., 2023). Considering these advantages and disadvantages, it is essential to examine the readiness of university students to apply AI to assist in writing academic essays.

Although the use of AI in academic writing has attracted substantial literature, empirical studies that investigate AI-supported writing practices in the context of Vietnamese EFL higher education remain limited. Furthermore, existing studies on attitudes and AI usage have not been able to provide multidimensional frameworks for comprehending learner readiness in academic writing with AI tools. Recent studies mainly concentrate on a single aspect of AI application, such as the level of technological competence or perceptions (González-Torres et al., 2023; Rafida et al., 2024; Sustaningrum & Haldaka, 2025). However, there is a lack of studies examining readiness in relation to cognitive, metacognitive, linguistic, ethical, and technological factors.

This study investigates the readiness of English-major undergraduates to engage in AI-supported academic writing. It also aims to explore the interrelationship between various readiness elements, such as cognitive-metacognitive, linguistic, ethical, and technology readiness. Moreover, this study examines the predictive effects of these readiness factors on the intention of AI-supported writing tools among undergraduate English majors.

To successfully achieve the study's purposes, the following main research questions will be addressed:

RQ1. To what extent are English-major undergraduates ready for AI-supported academic writing in terms of cognitive-metacognitive, linguistic, ethical, and technological dimensions?

RQ2. What are the relationships between English-major undergraduates' intention to use AI tools for academic writing and their cognitive–metacognitive readiness, linguistic readiness, ethical and responsible awareness, and technology readiness?

RQ3. To what extent do cognitive–metacognitive readiness, linguistic readiness, ethical and responsible awareness, and technology readiness predict students' intention to use AI tools?

2. Literature Review

2.1. AI-Supported Academic Writing in EFL Contexts

Artificial intelligence (AI) integration into academic writing has grown rapidly in the English as a Foreign Language (EFL) settings. According to Jasim et al. (2024) and Thi et al. (2025), AI applications enhance grammatical accuracy, structural coherence, and learning engagement as well as reduce writing anxiety.

Despite these advantages, AI use in academic writing also raises concerns. Alarcón et al. (2025) highlighted the risk of over-reliance. Ethical issues, including privacy, discrimination, joblessness, and security threats, have also been reported (Huang et al., 2023). Furthermore, AI tools may produce inaccurate suggestions due to limited understanding of meaning and context (Ngo et al., 2024).

2.2. Learner Readiness

One of the factors that contributes to the effective use of educational technologies is learner readiness (Almziad, 2025), which is regarded as a multidimensional construct. This implies that effective technology use requires not only technical capability but also cognitive competence, self-regulation, and awareness of ethical considerations. This multidimensional

perspective enables a better idea of the way students interact with AI-supported academic writing.

2.3. Dimensions of Readiness for AI-Supported Academic Writing

2.3.1. Cognitive–Metacognitive Readiness

The cognitive-metacognitive readiness is a notion that explains the ability of students to effectively use AI-facilitated technologies in academic writing through the implementation of cognitive and self-regulation strategies (Yurt, 2025). According to self-regulated learning theories, cognitive-metacognitive preparedness governs the manner in which learners plan, monitor and evaluate AI-supported writing (Teng & Zhang, 2016; Zimmerman, 2002). Within the framework proposed by Teng and Zhang (2016), cognitive readiness is characterized as Text Processing and Course Memory, and metacognitive readiness as Idea Planning and Goal-oriented Monitoring and Evaluation. The AI usage in writing is controlled by cognitive-metacognitive readiness, but ethical or technology readiness may control the attitudes or access (Ng et al., 2024). In this way, students who are more cognitively-metacognitively prepared are likely to implement more purposeful, strategic, and ethical use of AI-assisted writing tools.

2.3.2. Linguistic Readiness

Linguistic readiness can be defined as the ability of learners to utilize the language resources adequately (Ruminar & Gayatri, 2024). Linguistic readiness in AI-assisted writing is imperative because learners need to actively analyze and improve AI-produced texts instead of passively receiving them (Cui, 2025).

Linguistic readiness in the present study is conceptualised as learners' perceived capability to perform academic writing tasks in English. Rather than representing objective linguistic competence, this dimension is operationalised through writing self-efficacy, which reflects learners' beliefs about their ability to generate ideas, organise content, and apply appropriate language resources in writing (Limpo & Alves, 2017).

2.3.3. Ethical and Responsible Awareness

According to Pramod (2025), ethical awareness can be understood as the cognitive awareness of the consequences of using AI tools in decisions related to society and justice. In academic writing, ethical awareness involves issues of academic integrity, plagiarism, authorship, and responsible usage of AI-generated content (Kobrossy et al., 2025)

Based on recent research concerning AI literacy, ethical awareness does not only involve understanding AI regulations but also responsible AI use in educational contexts. Ng et al. (2023) suggested learners should be equipped with the ability to critically reflect on the ethical use of AI tools, which includes attribution, originality and responsible writing.

2.3.4. Technology Readiness

Technology readiness refers to an individual's disposition toward context-specific behaviors, namely, using new technologies for accomplishing goals (Parasuraman & Colby, 2015). The Tech Readiness Index (TRI) is based on this approach and is considered a gestalt of four personality scales: optimism and innovativeness are enablers, while discomfort and insecurity are inhibitors (Godoe & Johansen, 2012; Parasuraman, 2000).

Students who are more optimistic and innovative will actively apply various AI tools, but less comfortable and less secure students might be reluctant to use them successfully

(Marimon et al., 2025). Therefore, technology readiness is important in shaping learners' intention to use AI when writing (Rahman et al., 2025).

2.4. Intention to Use AI-Supported Writing Tools

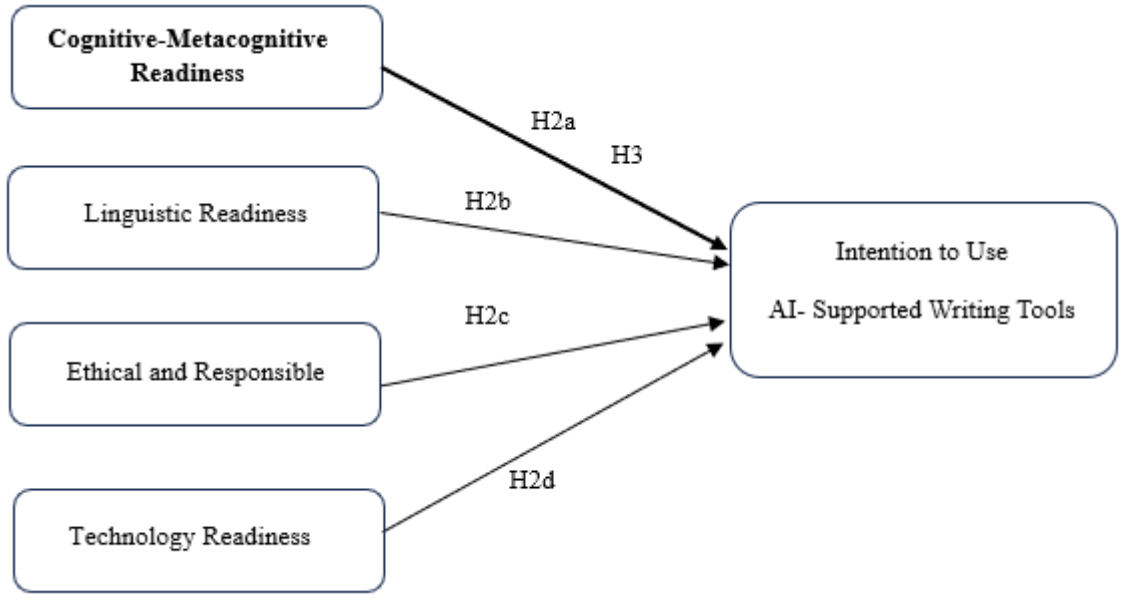
The Technology Acceptance Model (TAM) posits that users' intention is a direct predictor of actual technology use (Davis, 1989). Within the realm of AI-based academic writing, intention to use represents students' engagement in planning how they integrate AI tools, rather than occasional use. Although TAM typically focuses on perceived usefulness and perceived ease of use, our study broadens this approach by considering intention to use as a result of learner readiness, which encompasses cognitive-metacognitive, linguistic, ethical and technological dimensions. Therefore, TAM provides a theoretical basis for the dependent variables, but not a holistic framework.

2.5. Conceptual Framework

Cognitive–metacognitive readiness was adapted from the self-regulated learning strategies proposed by Teng and Zhang (2016). Linguistic readiness drew on four components of the English Writing Self-Efficacy scales by Bi and Wang (2025). Ethical and responsible awareness was informed by Ng et al. (2023), while technology readiness was adapted from Cui (2025) and the Technology Readiness Index model (Parasuraman & Colby, 2015). The outcome variable, intention to use AI, was adapted from the Technology Acceptance Model (Davis, 1989). This integrated framework was developed to provide a holistic assessment of undergraduates' readiness within the context of English language studies, while situating intention to use as an outcome construct rather than a full acceptance framework (see Figure 1).

Figure 1

Conceptual Framework of Learner Readiness and Intention to Use AI-Supported Writing Tools



Note. H3 refers to the comparative predictive strength among the four readiness dimensions.

Based on this conceptual framework, this study proposes the following hypotheses:

H1. The four readiness dimensions are expected to be positively associated with students' intention to use AI-supported academic writing tools.

H2a. Cognitive–metacognitive readiness significantly predicts students' intention to use AI-supported academic writing tools.

H2b. Linguistic readiness significantly predicts students' intention to use AI-supported academic writing tools.

H2c. Ethical and responsible awareness significantly predicts students' intention to use AI-supported academic writing tools.

H2d. Technology readiness significantly predicts students' intention to use AI-supported academic writing tools.

H3. Among the four readiness dimensions, cognitive–metacognitive readiness is hypothesized to exert the strongest predictive effect on students' intention to use AI-supported academic writing tools.

3. Methodology

3.1. Research Design

This study employs a mixed-methods approach, incorporating both quantitative and qualitative methods. Quantitative data were collected through a structured questionnaire. Qualitative data were obtained from semi-structured interviews.

3.2. Participants

The participants comprised 93 English-major undergraduate students at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT) (see Table 1).

Table 1

Participant Demographics (N = 93)

Variable	Category	n	%
Gender	Male	67	72
	Female	26	28
English proficiency	Intermediate	57	61.3
	Upper-intermediate	33	35.5
	Advanced	3	3.2
Frequently used AI tools	Generative AI tools (e.g., ChatGPT, Gemini, Claude)	88	94.6
	Grammar checking tools (e.g., Grammarly, ProWritingAid)	54	58.1
	Paraphrasing tools (e.g., QuillBot)	43	46.2
	Translation tools (e.g., DeepL/Google Translate)	29	31.2

Note. Percentages for AI tools exceed 100% because participants could select multiple responses.

All respondents were third-year students who had completed academic writing courses at the time of data collection.

Participants in the study were selected using a convenient sampling technique. Participation was voluntary, and all participants provided informed consent before data collection. Regarding prior experience with AI tools, most participants reported familiarity with Generative AI and grammar checkers.

3.3. Research Instruments

Two research instruments were employed: a questionnaire and semi-structured interviews.

The questionnaire included approximately 29 items measured on a five-point Likert scale ranging from strongly disagree to strongly agree. It was organized into seven sections: (1) introduction, (2) cognitive–metacognitive readiness, (3) linguistic readiness, (4) ethical & responsible AI awareness, (5) technology readiness, (6) intention to use AI in academic writing, (7) demographic information and AI usage background. To ensure linguistic accuracy and comprehensibility, the questionnaire was administered bilingually in English and Vietnamese using the back-translation technique. The overview of questionnaire dimensions, sources, and sample items is provided in Appendix A. In addition, five semi-structured interviews were also conducted to further explore students' experiences and perceptions of using AI in academic writing.

3.4. Data Collection Procedure

The online survey was deployed using Google Forms. Students also had easy access to the questionnaire since it was distributed through a QR code and could submit their responses immediately. After the data collection period, 93 entries were retrieved and uploaded to SPSS Statistics Version 26 software for quantitative analysis. Additionally, five semi-structured interviews were conducted through Zalo calls and recorded using the Audacity application. Each interview lasted approximately 20 minutes.

3.5. Data Analysis

Survey responses were obtained and summarized using descriptive statistics (mean and standard deviation). Moreover, Pearson correlation analysis was employed to test the relationship between variables, and the multiple regression analysis was conducted to determine predictors of students' intention to use AI-supported writing tools.

Qualitative data were analyzed using thematic analysis. Interview transcripts were coded and organized into themes in order to define the general trends and gain a more in-depth picture of the students' perceptions. Due to the anonymity of the participants, the interviewed students are referred to as S1 (Student 1), S2 (Student 2), and so forth.

3.6. Reliability and Validity

To ensure internal consistency, Cronbach's alpha coefficients were calculated for all sub-constructs in the questionnaire (see Appendix B). The results indicated acceptable reliability across the scales, with Cronbach's alpha values exceeding the threshold ($\alpha > .70$) for most dimensions. Following the Technology Readiness Index framework, technology readiness was examined as two distinct subdimensions: Technology Enablers and Technology Inhibitors. While Technology Enablers demonstrated satisfactory internal consistency ($\alpha > .70$), the reliability coefficient for Technology Inhibitors was slightly below the recommended threshold ($\alpha = .69$) but was still considered acceptable for an exploratory subscale (Hair et al., 2010; Taber, 2018).

Regarding validity, all measurement items were adapted from validated instruments in relevant studies. The questionnaire was administered bilingually in English and Vietnamese.

4. Results

4.1. Quantitative Results

4.1.1. Descriptive Statistics

The descriptive analysis was employed to analyse the degree of readiness of English-major undergraduates to write academic papers with the assistance of AI, based on four dimensions: (1) Cognitive - Metacognitive Readiness, (2) Linguistic Readiness, (3) Ethical and Responsible Awareness, and (4) Technology Readiness. Additionally, Technology Readiness has been categorised into two sub-constructs, Technology Enablers and Technology Inhibitors, to be explored deeply as discussed in the Methodology.

Table 2

Descriptive Statistics of Readiness Dimensions (N=93)

Readiness Dimensions	Mean (M)	Std. Deviation (SD)
Cognitive-Metacognitive	4.169	.570
Technology Enablers	4.149	.069
Ethical Awareness	4.141	.628
Technology Inhibitors	3.844	.893
Linguistic Readiness	3.688	.732
Intention to Use	4.096	.613

Note. Interpretation of Mean scores is based on the interval scale suggested by Pimentel (2010): 1.00–1.80 (Very Low); 1.81–2.60 (Low); 2.61–3.40 (Moderate); 3.41–4.20 (High); 4.21–5.00 (Very High).

The results in Table 2 revealed that the mean scores for all readiness dimensions ranged from 3.688 to 4.169. Cognitive-Metacognitive Readiness obtained the highest mean score (M = 4.169, SD = .570), although only marginally higher than Ethical Awareness (M = 4.141, SD = .628) and Technology Enablers (M = 4.149, SD = .069). Technology Inhibitors recorded a mean score of 3.844. The lowest mean score among the four dimensions was observed in Linguistic Readiness (M = 3.688, SD = .732). Regarding the dependent variable, the mean score for Intention to Use AI tools in academic writing was 4.096.

4.1.2. Correlation Analysis

The Pearson correlation results (see Table 3) indicated that all readiness dimensions, except Technology Inhibitors, had significant positive relationships with English major students' intention to use AI tools.

The strongest correlation was found between Technology Enablers and Intention (r = .635, p < .001), followed by Cognitive–metacognitive Readiness (r = .525, p < .001), Linguistic Readiness (r = .450, p < .001), and Ethical Awareness (r = .448, p < .001). Notably, Technology Inhibitors did not show a statistically significant correlation with intention (r = .167, p = .111).

Table 3

Pearson Correlation Matrix between Readiness Dimensions and Intention to Use AI (N=93)

Variables	1	2	3	4	5	6
1. Cognitive-Metacognitive	1					
2. Linguistic Readiness	.352**	1				
3. Ethical Awareness	.419**	.374**	1			
4. Technology Enablers	.552**	.483**	.512**	1		
5. Technology Inhibitors	.227*	.292**	.297**	.298**	1	
6. Intention to Use	.525**	.450**	.448**	.635**	.167	1
Sig. (2-tailed)	.000	.000	.000	.000	.111	

Note. p = Pearson's correlation coefficient; **(2-tailed): p-values < .01, *(2-tailed): p-values < .05

4.1.3. Multiple Regression Analysis

A multiple regression was conducted to identify the strongest predictors of intention. The model was statistically significant, $F(5, 87) = 16.064$, $p < .001$, explaining 48% of the variance in students' intention ($R^2 = .480$).

As shown in Table 4, Technology Enablers ($\beta = .404$, $p < .001$) and Cognitive-Metacognitive Readiness ($\beta = .216$, $p = .025$) were the only two significant predictors. Other factors, including Linguistic Readiness, Ethical Awareness, and Technology Inhibitors, did not substantially contribute to the model ($p > .05$). All VIF values were under 2.0, indicating that there are no issues related to multicollinearity.

Table 4

Multiple Regression Analysis Predicting Intention to Use AI (N=93)

Predictors	β	t	p	VIF
Cognitive-Metacognitive	.216	2.28	.025	1.505
Linguistic Readiness	.16	1.762	.082	1.383
Ethical Awareness	.115	1.23	.222	1.469
Technology: Enablers	.404	3.841	.000	1.851
Technology: Inhibitors	-.084	-1.007	.317	1.161

Note. Model Statistics: $R = .693$; $R^2 = .480$; Adjusted $R^2 = .450$; $F = 16.052$; $p < .001$.

4.2. Qualitative results

4.2.1. Cognitive–Metacognitive Readiness in AI-Supported Writing

All participants demonstrated clear awareness of the AI's limitations and mentioned critical thinking as an indispensable part of their writing process. However, clear differences emerged across proficiency levels. Upper-intermediate students (S1 and S5) exhibited strong metacognitive readiness through initiating questioning of the output reliability and establishing personal benchmarks.

"I use AI to assist me with brainstorming ideas. Right from the start, … I need to refine that topic to see if it's practically suitable, as well as get ideas on what exactly I should cover within that topic." (S1)

"Typically, I'm confident with a 70-30 ratio, where 70% comes from personal effort, and the remaining 30% comes from AI." (S5)

Conversely, intermediate students (S2 and S3) were more inclined to interact with AI tools through direct experience instead of anticipatory planning. Both S2 and S3 described the cases when AI provided them with inaccurate information and responded by cross-referencing with Google or other AI platforms.

4.2.2. Linguistic Readiness for AI Usage

The results of the interview illustrated the greatest distinction between the two proficiency groups. The upper-intermediate students emphasised that they primarily utilized AI to polish their expression and enhance academic writing skills, whereas the intermediate-proficiency students considered AI as one of the means to compensate for their language shortcomings.

"I could say I have to use AI to ask it things like, "How can I write this more formally?", and it guides me in several ways … AI is just there to support and give suggestions." (S1)

Specifically, S2 conveyed feelings of insecurity when comparing their own work to AI-generated texts and uncertainty about their ability to achieve the required score without that assistance.

"Sometimes, I feel like I'm becoming too dependent on this AI. And without it, I don't know how long it would take me to achieve my target score." (S2)

4.2.3. Ethical and Responsible Awareness in AI Use

All participants set boundaries on their AI usage, but their self-regulation mechanisms differed considerably. S1 made a clear distinction between "the application of AI tools to support" and "copy-paste". Additionally, S1 added "human perspectives" to prevent a mechanical writing style. S2 established a technical indicator of AI under 30 per cent. S3 used metaphors, the pillars and the roof, to illustrate a real sense of the border between support and replacement.

"The principle I always apply is that AI provides the foundation for my work. I just need to step up and build the pillars and roof myself to complete my part." (S3)

S5 applied the principle of using AI for one section, then independently wrote the following sections, demonstrating a remarkable finding regarding the structured self-regulation mechanism and practical applications of AI in academic writing.

4.2.4. Technology Readiness

All participants were in a positive state and ready to experiment with new tools. The approach to new tools was consistent, including trying with simple questions first, comparing with the old tools, and then applying it to the actual task (S2, S3, and S4). S1 focused on paid work and "training" AI according to individual needs.

"I have to train and evaluate AI's authenticity through data, difficult problems or things I pose for it to solve." (S1)

While S2 was the only case showing genuine concern about how much time they needed to reach their expected score without the AI support, the other participants perceived the risk of dependence but were not overly anxious. By contrast, neither upper-intermediate English learner was worried about AI disappearing.

5. Discussion

5.1. Students' Readiness for AI-Supported Academic Writing

Notably, the highest mean score (M = 4.17) of Cognitive-Metacognitive Readiness indicates that participants have self-regulated learning abilities, which entail the process of planning, monitoring, and evaluating AI-generated content. This result is consistent with the study of Tezer (2025), which showed that AI systems promote metacognitive engagement, an essential process that allows learners to acquire knowledge actively and adjust their learning methods strategically.

Simultaneously, the relatively high scores for Technology Enablers and Ethical Awareness dimensions suggested that students' intention to use AI may be associated not only with access to technology but also with responsible engagement with AI tools. This finding represents a broader understanding of AI literacy among EFL undergraduates, beyond technical proficiency, entailing the scope of evaluation, metacognition, and ethical responsibility in AI-supported academic writing practices (Adamakis & Rachiotis, 2025).

However, the mean score of the Technology Inhibitors was relatively low, suggesting that a group of learners still experienced digital anxiety and insecurity. This finding aligns with a study by Almusharraf (2026), which shows that learner anxiety is a relevant factor in the context of AI-assisted writing, where perceived usefulness is coupled with increased anxiety when using AI chatbots for writing in English as a second language.

The difference in mean values between Cognitive-Metacognitive Readiness and Linguistic Readiness can be interpreted through the self-regulated learning theory (Zimmerman, 2002), which assumes that learners become more sensitive to their linguistic constraints as they become more strategically aware. Recent studies indicate that students with an advanced metacognitive capacity are more readily engaged in the process of critical evaluation, revision, and strategic adaptation; thus, they are also more likely to become aware of the gaps in their linguistic accuracy and sensitivity (Sun & Wang, 2020; Teng, 2025).

The variations in AI readiness dimensions were further illuminated by qualitative data from participants with different levels of English proficiency. Even though all interviewees were conscious of the constraints of AI tools and the significance of critical thinking, upper-intermediate-English-level students displayed a greater level of self-management through the process of developing outlines, testing reliability and setting personal benchmarks. The intermediate-English-level students, in turn, were more passive in their attitude towards AI, as they mostly verified AI-generated texts only when they suddenly detected a potential error.

5.2. Relationships among Readiness Dimensions and Intention to Use AI

The results of the Pearson correlation analysis largely confirm Hypothesis H1, indicating that most aspects of readiness examined in this study are positively correlated with students' intention to use AI tools in academic writing. Specifically, the strongest bivariate association between Technology Enablers and Intention to Use implies that a positive and active approach toward technology identifies the willingness of students to adopt AI tools. This finding is aligned with the study of Dahri et al. (2024), who propose that the enabling conditions and the perceived self-efficacy within the AI-interacting environment are crucial determinants of AI adoption.

The fact that there was no statistically significant correlation between Technology Inhibitors and Intention to Use is further clarified by qualitative data. In other words, technological barriers do not seem sufficient to influence English major undergraduates'

intention to use AI. For upper intermediate-level students (S1, S5), AI tools play a more supportive than essential role; as a result, these students expressed little concern if AI becomes unavailable. Conversely, S2, an intermediate-level student, was particularly concerned that his grades might be affected without the support of AI tools.

Moreover, the relationship between Intention to Use and Ethical Awareness was further contextualized by the qualitative results. In particular, the interviewees demonstrated multiple attitudes towards the ethical responsibility of AI usage, such as balancing human-AI cooperation and complying with the principle of less than 30% AI content.

5.3. Predictive Power of Readiness Dimensions

The regression analysis results showed a differentiated pattern when considering the predictive power of readiness dimensions on EFL undergraduates' intention to adopt AI tools for academic writing. The regression coefficient of Cognitive-Metacognitive Readiness showed significant influence on Intention to Use, thus confirming Hypothesis H2a. Nevertheless, contrary to Hypothesis H3, which suggested that Cognitive-Metacognitive Readiness is the strongest predictor of Intention to Use, the regression results demonstrated that Technology Enablers had a more substantial predictive effect.

On the other hand, Technology Inhibitors did not have a statistically significant effect on Intention to Use, which implies that the sentiment of discomfort or insecurity can be alleviated by perceived benefits and increased familiarity with AI tools. Based on the current findings, optimism and willingness to experiment with AI appeared to be driving factors for technology use; this is in line with the Technology Readiness Index (Parasuraman, 2000), which suggests that a lack of anxiety does not necessarily mean that learners do not perceive barriers, such as concerns about reliability or over-reliance on AI. Additionally, these results correspond to a recent study that highlights the crucial role of positive beliefs about technology in influencing the intention to use AI (Sustangingrum & Haldaka, 2025).

In opposition to hypotheses H2b and H2c, Linguistic Readiness and Ethical Awareness were not significant predictors of Intention to Use in the multivariate model, even though both dimensions were significantly correlated with Intention to Use in the bivariate correlation analysis.

Taken together, these findings suggest that cognitive-metacognitive and ethical preparation play an essential role in guiding how AI is used, rather than directly determining students' intention to adopt AI. Nevertheless, EFL undergraduates' intention to apply AI in academic writing appears to be largely dependent on their attitudes towards technological support, previous positive experiences, and self-regulation in the use of AI tools.

6. Conclusion

In general, the descriptive results showed that the participants reported relatively high levels of readiness across the examined dimensions. Meanwhile, the findings of correlation analysis revealed that the majority of readiness dimensions were positively associated with the intention to use AI. However, the regression analysis results unveiled that the cognitive-metacognitive aspect did not exhibit the strongest predictive power. The technological enabling factor was found to have the highest predictive coefficient on the intention to use AI. On the other hand, linguistic readiness and ethical awareness, which are the critical factors of responsible use, did not significantly influence the adoption decisions.

Based on qualitative interview results, the study found that technological competence and positive attitude are the driving factors behind the intention of EFL students to utilize AI

in writing academic papers, yet cognitive and ethical preparation function as moderating and controlling factors in how AI is used rather than whether it is adopted.

7. Implications and Limitations

7.1. Theoretical Implications

The present study offers preliminary evidence that AI readiness in TESOL is a multidimensional construct, which needs to distinguish between cognitive-metacognitive readiness and technological support in predicting the intention of learners to use AI. Although metacognitive preparedness is the guarantee of the quality and integrity of AI-supported writing, positive beliefs and supporting technological aspects are the main factors that can influence the choice of learners to adopt AI, and future theoretical models should consider both the ability to exert influence and the support conditions offered by technology.

7.2. Pedagogical Implications

Given that students' intention to adopt AI is primarily driven by positive beliefs toward technological affordances, English writing instructors should design learning tasks that allow students to experience, compare, and evaluate multiple AI tools in academic writing contexts. To prevent overreliance, AI should be positioned as a controlled support mechanism at specific stages of the writing process, such as outlining, checking coherence, or editing. These tasks should be accompanied by structured reflection requirements, including short reflective writing (approximately 150–200 words) or an AI Use Declaration. Through these practices, learners are encouraged to regulate their AI usage and align technological affordances with ethical judgment. On an institutional level, universities should offer consistent directions on how AI can be utilized in academic writing and provide students with access to credible AI platforms.

7.3. Limitations and Future Research

Despite the significance of the contributions to the study, there are still limitations that need to be considered. Firstly, the sample size and the scope of the institution are limited, as only EFL students of one institution of higher learning were used. This may affect the generalizability to other learning contexts. Second, the cross-sectional design fails to provide the opportunity to monitor the transformation of the perceptions, attitudes, and behavior of learners towards the use of AI over time.

Consequently, future studies should extend to longitudinal studies to examine how AI readiness and intentions to use AI evolve during a course or training program. Moreover, empirical or quasi-experimental research may be performed to evaluate the effects of various models of AI integration in the teaching of academic writing.

Acknowledgments

The authors would like to express sincere gratitude to third-year students majoring in Business English, who voluntarily took part in this study. Special thanks are extended to Mr. Tran Thanh Tu, the lecturer of the Introduction to Business subject at HUFLIT, for allowing the survey to be conducted in his classes and for his constructive feedback.

References

Adamakis, M., & Rachiotis, T. (2025). Artificial intelligence in higher education: A state-of-the-art overview of pedagogical integrity, artificial intelligence literacy, and policy integration. *Encyclopedia, 5*(4), 180. https://doi.org/10.3390/encyclopedia5040180

Akanda, F., & Talukder, T. (2026). AI writing tools in English writing: Reviewing teachers' and students' views on benefits and concerns. *Social Sciences & Humanities Open, 13*. https://doi.org/10.1016/j.ssaho.2026.102437

Alarcón, C. M. M., Lasekan, O. A., & Pena, M. T. G. (2025). AI concerns among educators: A thematic analysis. *International Research Journal of Multidisciplinary Scope, 6*(1), 1031–1045. https://doi.org/10.47857/irjms.2025.v06i01.01929

Almusharraf, A. (2026). AI chatbots in L2 writing: Anxiety, usefulness, and writing outcome. *Humanities and Social Sciences Communications, 13,* 393. https://doi.org/10.1057/s41599-026-06816-w

Cui, Y. (2025). What influences college students using AI for academic writing? – A quantitative analysis based on HISAM and TRI theory. *Computers and Education: Artificial Intelligence, 8*, 100391. https://doi.org/10.1016/j.caeai.2025.100391

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon, 10*(8), e29317. https://doi.org/10.1016/j.heliyon.2024.e29317

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *Frontiers in Public Health, 11*, 1166120 https://doi.org/10.3389/fpubh.2023.1166120

Godoe, P., & Johansen, T. S. (2012). Understanding adoption of new technologies: Technology readiness and technology acceptance as an integrated concept. *Journal of European Psychology Students, 3*(1), 38–52. https://doi.org/10.5334/jeps.aq

González-Torres, P., Cabrera-Solano, P., & Castillo-Cuesta, L. (2023). Stakeholders' perceptions of teaching and technological skills in EFL vocabulary instruction: Implications for remote learning. International Journal of Learning, *Teaching and Educational Research, 22*(7), 173–192. https://doi.org/10.26803/ijlter.22.7.10

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Pearson.

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence, 4*(4), 799–819. https://doi.org/10.1109/TAI.2022.3194503

Jasim, M. Y., Musa, Z. H., Asim, Z. A., & Salman, A. R. (2024). Developing EFL writing with AI: Balancing benefits and challenges. *Technology Assisted Language Education, 2*(2), 80-93. https://doi.org/10.22126/tale.2024.10953.1052

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*. https://doi.org/10.1016/j.cmpbup.2024.100145

Kobrossy, S., Jasim, Y., Ismail, I. A., Goud, N. N., Jasim, S., & Awqati, A. (2025). Artificial intelligence and the development of academic authorship: Analyzing AI-generated content, ethical issues, and intellectual property challenges. *International Journal of Management and Data Analytics, 5*(1), 253–270. https://ijmada.com/index.php/ijmada/article/view/99

Limpo, T., & Alves, R. A. (2017). Relating beliefs in writing skill malleability to writing performance: The mediating role of achievement goals and self-efficacy. *Journal of Writing Research, 9*(2), 97-125. https://doi.org/10.17239/jowr-2017.09.02.01

Liu, C., Hou, J., Tu, Y.-F., Wang, Y., & Hwang, G.-J. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments, 31*(9), 5614–5632. https://doi.org/10.1080/10494820.2021.2012812

Marimon, F., Arias Valle, M. B., Coria Augusto, C. J., & Larrea Arnau, C. M. (2025). From optimism to trust: How ChatGPT is reshaping student confidence in AI-driven learning. *RIED-Revista Iberoamericana de Educación a Distancia, 28(2),* 131–153. https://doi.org/10.5944/ried.28.2.43238

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2023). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology, 55*(3), 1082-1104. https://doi.org/10.1111/bjet.13411

Ngo, A., Gupta, S., Perrine, O., Reddy, R., Ershadi, S., & Remick, D. (2024). ChatGPT 3.5 fails to write appropriate multiple choice practice exam questions. *Academic Pathology, 11*(1), Article 100099. https://doi.org/10.1016/j.acpath.2023.100099

Parasuraman, A. (2000). Technology readiness index (TRI): A multiple-item scale to measure readiness to embrace new technologies. Journal of Service Research, 2(4), 307–320. https://doi.org/10.1177/109467050024001

Parasuraman, A., & Colby, C. L. (2015). An updated and streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research, 18*(1), 59–74. https://doi.org/10.1177/1094670514539730

Pramod, D. (2025). Decoding responsible AI use: The influence of digital literacy and ethical awareness. *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2025.2592371

Rafida, T., Suwandi, S., & Ananda, R. (2024). EFL Students' perception in Indonesia and Taiwan on using artificial intelligence to enhance writing skills. *Jurnal Ilmiah Peuradeun, 12*(3). https://doi.org/10.26811/peuradeun.v12i3.1520

Rahman, M. K., Hossain, M. A., Ismail, N. A., Hossen, M. S., & Sultana, M. (2025). Determinants of students' adoption of AI chatbots in higher education: The moderating role of tech readiness. *Interactive Technology and Smart Education, 22*(3), 448–480. https://doi.org/10.1108/ITSE-12-2024-0312

Ruminar, H., & Gayatri, P. (2024). Linguistic readiness of doctoral candidates: A case study of TOEFL policy at the Faculty of Agriculture. E-Structural, 7(2), 78–94. https://doi.org/10.33633/es.v7i02.10933

Setiana, L. N., & Chamalah, E. (2025). Revolutionizing scientific writing: The role of AI in enhancing efficiency, accuracy, and creativity. *Journal of Language, Literature, and Cultural Dynamics, 2*(01), 40-47. https://doi.org/10.65840/jllcd.v2i01.30

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System, 90*. https://doi.org/10.1016/j.system.2020.102221

Sustaningrum, R., & Haldaka, M. (2025). Student utilization and perceptions of AI technology for academic purposes: *A mixed-method analysis. Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2025.2553835

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in Science Education, 48(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Teng, L. S., & Zhang, L. J. (2016). A questionnaire-based validation of multidimensional models of self-regulated learning strategies. *Modern Language Journal, 100*(3), 674–701. https://doi.org/10.1111/modl.12339

Teng, M. F. (2025). Understanding EFL student writers' metacognitive awareness in utilizing ChatGPT. *System, 135,* 103848. https://doi.org/10.1016/j.system.2025.103848

Tezer, M. (2025). Metacognitive engagement in AI-supported learning: Frameworks, challenges, and transformations. In M. Tezer (Ed.), *Metacognition in action: Insights into learning, AI, and self-regulation*. IntechOpen. https://doi.org/10.5772/intechopen.1012658

Thi, X. H. N., Thien, H. V. H., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through AI-powered tools: perceived benefits and challenges among Vietnamese EFL students. *Discover Education, 4*(1). https://doi.org/10.1007/s44217-025-00905-9

Yurt, E. (2025). The self-regulation for AI-based learning scale: Psychometric properties and validation. *International Journal of Current Educational Studies, 4*(1), 95–118. https://doi.org/10.46328/ijces.176

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Bionote

Tran Phuong Anh is an EFL teacher and researcher. She holds a Master's degree in English Linguistics from Ho Chi Minh City University of Foreign Languages and Information Technology. She is interested in academic essay writing and AI-assisted learning. She is dedicated to developing learner autonomy in English classrooms. She aspires to pursue a PhD in TESOL and Applied Linguistics, with a concentration in academic literacy and technology-enhanced language learning.

Ngo Thi Thu Hien is a master's graduate in English Language from Ho Chi Minh City University of Foreign Languages and Information Technology. She is now teaching online from home, including the Cambridge English program. She helps students build language competence and prepare for exams. She has also delivered a conference paper, advancing

research in English language teaching methods. Her research includes language teaching, motivation, and classroom techniques.

Appendix A

This appendix presents an overview of the questionnaire dimensions used in the study.

Table A1

Questionnaire Dimensions, Sources, and Sample Items

Dimension	**Theoretical Source(s)**	**No. of Items**	**Sample Item**
Cognitive–Metacognitive Readiness	Zimmerman (2002); Teng and Zhang (2016)	7	*Before writing, I use AI to search for related ideas and information to help me plan my composition.*
Linguistic Readiness	Bi and Wang (2025)	7	*I can organize AI-generated sentences into a logical and cohesive academic paragraph.*
Ethical Awareness	Ng et al. (2023);	5	*I evaluate whether the AI-generated information is derived from reliable and credible sources.*
Technology Enablers	Parasuraman and Colby (2015)	4	*I actively keep up with the latest AI developments to enhance my English learning and writing.*
Technology Inhibitors	Parasuraman and Colby (2015)	2	*I sometimes feel that AI tools are not helpful because they cannot explain writing concepts in a way I can understand.*
Intention to Use AI	Davis (1989)	4	*I intend to use AI tools to support my English academic writing in the next semester.*

Note. Table A1 summarizes the adapted theoretical source for each dimension, the number of items, and a typical sample survey question for each dimension. The full questionnaire is available upon request.

Appendix B

This appendix reports the reliability of the dimensions.

Table B1

The Reliability of Questionnaire Constructs

Construct	**No. of Items**	**Cronbach's Alpha (α)**
Cognitive–Metacognitive Readiness	7	.811
Linguistic Readiness	7	.853
Ethical Awareness	5	.783
Technology Enablers	4	.742
Technology Inhibitors	2	.690
Intention to Use	4	.757

Note. Cronbach's alpha values above .70 indicate acceptable internal consistency. The slightly lower alpha value for Technology Inhibitors is considered acceptable for an exploratory subscale with a limited number of items.

#41 Students' Perceptions of Google Meet in ESL-Oriented Online Cambridge Exam Preparation at Private Language Centers in Ho Chi Minh City

Ngô Thị Thu Hiền[1]

Abstract

As digital technology rapidly advances, online learning of English has become a new trend in numerous teaching environments. The purpose of this study is to examine how learners view ESL-focused online English learning using Google Meet in private classes of Cambridge exam preparation in Ho Chi Minh City. The study employed a mixed-method approach involving 35 KET and PET level students, utilizing semi-structured interviews and a Likert scale questionnaire analyzed through SPSS. Results indicated a positive perception of the online learning environment, with students appreciating communication with instructors and peers, and effectively using English. Google Meet features like screen sharing and chat were noted as beneficial for understanding and engagement. Limitations included a small sample size, suggesting future research should involve larger and more diverse groups to further explore digital technologies in online ESL learning.

Keywords: ICT, online ESL learning, Google Meet, digital tools

1. Introduction

In Vietnam, English as a foreign language (EFL) has always been taught with the classroom practices focusing more on grammar knowledge and test results than the application of meaningful use of language (Ha et al., 2021). Nevertheless, the growing academic, professional and international requirements towards the English language have contributed to the rising voices in favor of a change into an English as a Second Language (ESL) orientation, which is more communication-oriented, inter-personal and life-based in language use (Hsu, 2015). This shift has emerged as one of the main topics of recent debate about the English language education in Vietnam.

Meanwhile, online learning is now an element of the English language teaching, especially after the Covid-19 pandemic (Rintaningrum, 2023). Synchronous learning has also opened up new possibilities in interaction, sharing and exposure to the English language outside the classroom setting due to digital platforms (Procel et al., 2024). Google Meet is one of these platforms, as its availability, convenience, and integration with other computer technologies that facilitate online learning have become very popular (Al-Maroof, 2020).

The effectiveness of online applications in enhancing language skills and learning outcomes of students has been investigated in many studies such as studies by Al-Jarf (2026) and Popovici and Mironov (2015). Nevertheless, minimal studies have been conducted on the perceptions of ESL-based online teaching of English among learners in Vietnam in private classes of exam preparation in Cambridge English. The majority of past research is inclined to consider the general online learning settings or university setting instead of domestic language classes that prepare students to pass international English exams.

[1] *A graduate of the postgraduate program at Ho Chi Minh City University of Foreign Languages and Information Technology; Email: friendship200561@gmail.com*

Moreover, there are numerous tools offered by digital platforms like Google Meet that can support the interaction, communication, and engagement of online synchronous lessons (Eraković & Topalov, 2021). These tools could be successful in establishing an ESL-friendly learning atmosphere in which students are motivated to learn English actively during classroom practices. Nevertheless, the success of these tools is highly reliant on the perceptions and experiences of students of online learning.

The following research questions are addressed in the study:

RQ1: How do students perceive ESL-oriented online English learning through Google Meet in private classes of Cambridge exam preparation in Ho Chi Minh City?

RQ2: What is the perception of learners regarding digital tools utilization on Google Meet as an assistant to the processes of interaction, engagement, and communication in English concerning online ESL-oriented courses?

2. Literature Review

2.1 Information and Communication Technology (ICT) and development of e-learning

According to AlTameemy and Alrefaee (2021), ICT is becoming an increasingly important component of our daily lives, as it is a natural consequence of the significant advancements in modern technology. Adaptive e-learning is viewed as stimulation to support learning and improve student engagement, so designing appropriate adaptive e-learning environments contributes to personalizing instruction to reinforce learning outcomes (El-Sabagh, 2021). According to El-Seoud et al. (2014), E-learning has become increasingly important as a result of advancements in technology, significantly transforming higher education. The 21st century has witnessed rapid development in online learning, with enrolment increasing by 12-14 percent annually from 2004 to 2009. This growth is largely attributed to the flexibility e-learning offers, allowing students to access education from anywhere and at any time, with options for part-time or full-time study. Additionally, e-learning facilitates easier information sharing among students, enhancing educational experience.

2.2 EFL and ESL in English Language Education

According to Zhou et al., (2024), English language teaching has been long conceptualized in the context of the difference between English as Foreign Language (EFL) and English as Second Language (ESL). According to the research by Religioni et al (2024), ESL and EFL are differentiated by the context of learning and the level of language interaction. In an ESL setting, English exposure is more frequent and varied than in an EFL setting. More significantly, ESL offers greater chances of negotiating the meaning, as learners frequently request clarification and are presented with altered input by interlocutors. The frequency of adjustment and reformulation of the output is also promoted among ESL learners, which allows them to identify and address language issues. Sato and Storch (2020) explained that the difference between ESL and EFL settings is determined by learners' needs and purposes. On the cognitive and structural level, an ESL context facilitates language learning by offering more quality, quantity, and diversity of language input and multiple opportunities for use of the target language in both the classroom and out. Thus, ESL provides more productive conditions for acquiring a second language, compared to EFL.

2.3 Digital Tools and Synchronous Online English Learning

The quick growth of online education has revolutionized the education of English language, especially by adopting the synchronous digital platform (Moorhouse & Wong, 2021). Real-time communication between teachers and learners is possible with help of synchronous online learning environments providing an opportunity to use communicative language, to provide instant feedback and collaborate with teachers and other learners. A study by Radovan et al. (2020) integrates our experiences and comparative analysis of using different tools for online education during times where the usual teaching methods are not an option because of various restrictions related to travel bans, quarantines or virus pandemics. The tools being analyzed are Skype, Microsoft Teams, Adobe Connect and Zoom. Several aspects are analyzed including starting courses and creating groups of students, enrolling students to the class, delivering the lecture, communicating with the students during the class, getting feedback from the students in a virtual class, tracking student activity, creating live polls and questionnaires, sharing content, recoding audio and video content, submitting assignments and homework and accessing the archived content. Each explored aspect has a direct influence on the whole teaching and learning experience as well as the effectiveness of supporting the educational processes.

Google Meet is one of those that have gained popularity as a widely used platform because of its accessibility, ease of use and compatibility with other Google based applications (Al-Maroof et al., 2021). These characteristics enable educators to create engaging online courses that facilitate ESL-oriented practices even during exam preparation. Besides, the Google Meet platform is widely utilized for video conferencing in online learning, facilitating audiovisual communication between teachers and students in both small and large groups (Nehe, 2021). It offers various features that enhance online learning by allowing teachers to present materials and encourage interactive discussions (Purwoko et al., 2022). According to Yumnam (2021), Google Meet is accessible video conferencing platform suitable for both business meetings and online classes. These services allow users to log in via various devices. They facilitate live interaction between teachers and students, enabling the sharing of presentations and multimedia resources. In ESL classes, teachers can cover essential language skills using PowerPoint and interactive methods (Zahara & Jupri, 2022). While students can practice concepts through direct engagement with instructors, they are given the opportunity to practice the skills being learnt in real time and then are provided with scaffolding from their teacher that is empathetic and immediate, they move from being passive recipients to active problem solvers (Zhang, 2023).

2.4 Related Studies

Solihah (2023) discovered that Google Meet was an effective tool that facilitated online learning in a TOEFL preparation course by making the interaction between students and the lecturer real-time, boosting comprehension and motivation. The study by Pham and Nguyen (2023) shows that university students in Vietnam positively perceived the use of Google tools—Google Classroom, Google Forms, Google Meet, and Google Sites—for learning English. These tools enhanced their educational experience and improved their progress in English, highlighting the effectiveness of Google products in language instruction and suggesting avenues for further research in online English education technologies. The research by Yumnam (2021) established that e-learning is an efficient means of teaching ESL since it helps improve language of students, boosts their interaction and offers them flexibility in learning, but proper technological support and planning are needed to implement it effectually.

3. Methodology

3.1 Research Design

In order to obtain data to conduct this mixed-method study, a questionnaire and semi-structured interviews were used. The questionnaire itself was a set of Likert-scale items created with the help of Cognito Forms with 1-5 inherent to the statement of Strongly Disagree and Strongly Agree respectively. The questionnaire employed in the current study was adapted modified from Delviyanti et al. (2022) for the context of online Cambridge exam preparation classes. The method of questionnaire was chosen because it made it possible to gather systematized information about the opinions of the learners concerning the use of language, interaction, engagement, and the contribution of digital tools into the online learning settings. The research aimed at the description of the perceptions of the learners, not quantifying the learning results or causal relationships. Furthermore, semi-structured interviews were carried out with some volunteer students to gain deeper insights into their perception of online learning of the English language through Google Meet. The interviews were held with the help of Google Meet video calls and took about 25 minutes.

3.2 Participants

The respondents used in this research were 35 students who are taking privately online classes in Cambridge English examination preparation in Ho Chi Minh City in Vietnam. The learners were KET and PET level. Everyone possesses a minimum of 1 year's experience in utilizing Google Meet for learning purposes. The two levels of proficiency were chosen because the learners were deemed as mature and linguistically competent enough to appreciate the questionnaire items and give credible self-reported answers.

KET and PET learners were generally capable of sufficient reading comprehension and had previous exposure to online learning in English, which allowed them to reflect on the learning experience and attitudes towards ESL-oriented online learning. Younger learners in lower levels of Cambridge were, on the contrary, excluded because of age factors such as cognitive and linguistic constraints that could have compromised the validity of the questionnaire answers.

3.3 Data Collection Procedures

The questionnaire was conducted via Cognito Forms after the students had received ESL-focused English lesson via Google Meet. The participants had one week to answer the questionnaire, guaranteeing time for thoughtful consideration. The questionnaire was administered as an online survey, and students filled it out on their own in preparation time. All responses were received within the allotted seven-day time frame.

Following this step, five students were invited to take part in an online semi- structured interview through Google Meet. The interview took almost 25 minutes. The interviews were conducted and then transcribed and coded thematic interview protocols.

3.4 Data analysis

Data obtained from the questionnaires was analyzed in SPSS Version 26 (Statistical Package for the Social Sciences). All responses were screened for completeness and correct answers before analysis. Descriptive statistical analyses were to summarize the learners' perceptions, Mean, standard deviation and percentage. The interview data included content analysis, a process that involves collecting and interpreting data based on related concepts and topics. Quantitative data can be dealt with statistical analysis to help explain the patterns and frequency of the issue, while qualitative data from interviews provide profound insights.

3.5 Reliability

The Statistical Package for the Social Sciences (SPSS) version 26 was used to determine the Cronbach alpha in order to make sure that the questionnaire was reliable. The scores revealed that the constructs had acceptable reliability with alpha ranging between 0.879 and 0.912 (see Table 1). This indicates that the questionnaire questions were internally reliable and could be analyzed further.

Table 1

Reliability of the questionnaire and constructs

Constructs	N of items	Cronbach's Alpha value
Google Meet and Digital Tools	5	0.879
ESL-Oriented Language Use	5	0.891
Teaching Practices and Communication	5	0.897
Engagement and Overall Perception	5	0.907
Total	20	0.963

4. Findings

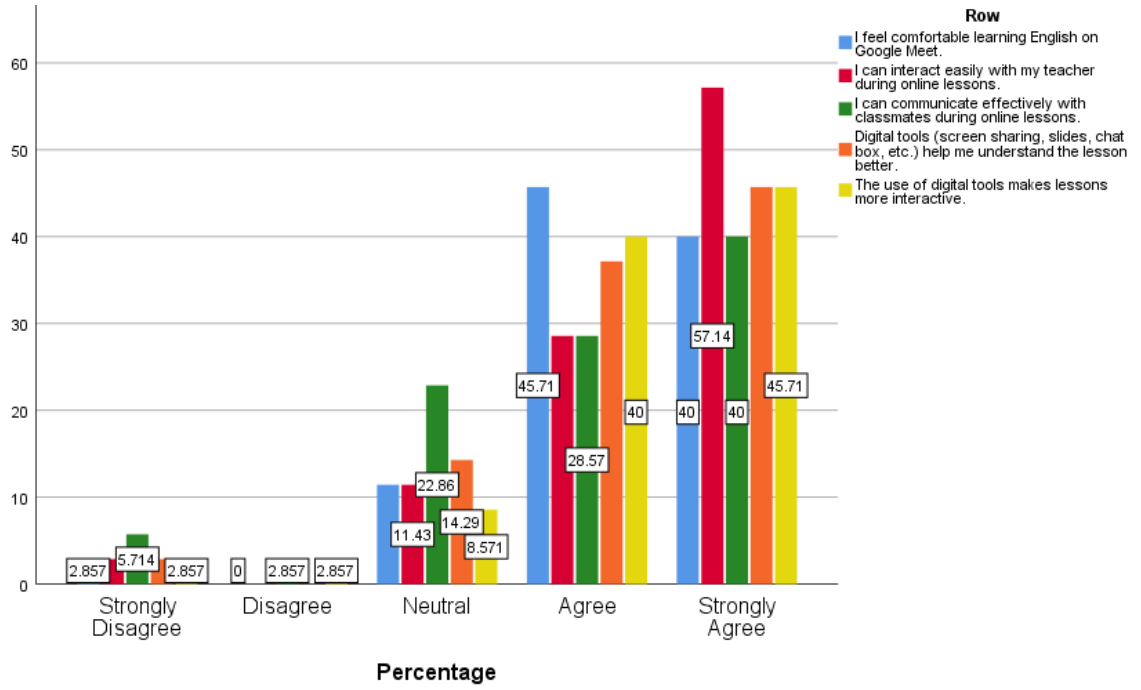
4.1 Findings from questionnaire

4.1.1 Google Meet and Digital Tools

The results shown in Figure 1 indicate that the perceptions of the students about learning English through Google Meet were very positive. There was a high percentage of students who said that they were at ease learning on the platform with 45.71% responding in the affirmative, and 40 percent strongly affirming this statement (M=4.20, S.D.=0.868). Regarding contact with the teacher, 28.57 and 57.14 percent agreed and strongly agreed that they could easily interact with the teacher during online lessons (M=4.37, S.D.=0.910). Likewise, when it comes to communication with classmates, 28.57% agreed and 40% communicate effectively in online classes (M=3.94, S.D.=1.136). Utilization of digital tools also received a positive response with 37.14 percent of the respondents agreeing to the tool and 45.71 percent strongly agreeing that the tool assisted them to understand the lesson better (M=4.23, S.D.=0.910). In addition, 40% and 45.71% responded that the lessons were more interactive with the help of digital tools and strongly agree, respectively (M=4.23, S.D.=0.942). Altogether, the agreement rates are quite high, which proves that Google Meet and its online resources are effective to support the learning of English online among students.

Figure 1

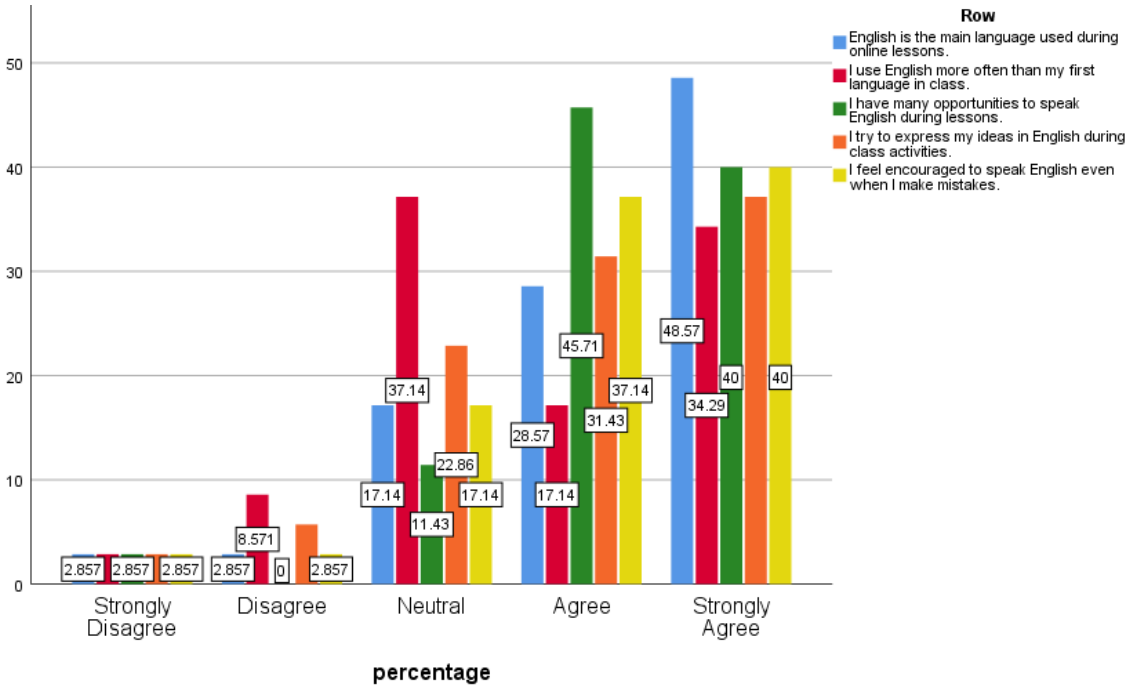
Google Meet and Digital Tools



4.1.2 ESL-Oriented Language Use

Figure 2.

ESL-Oriented Language Use



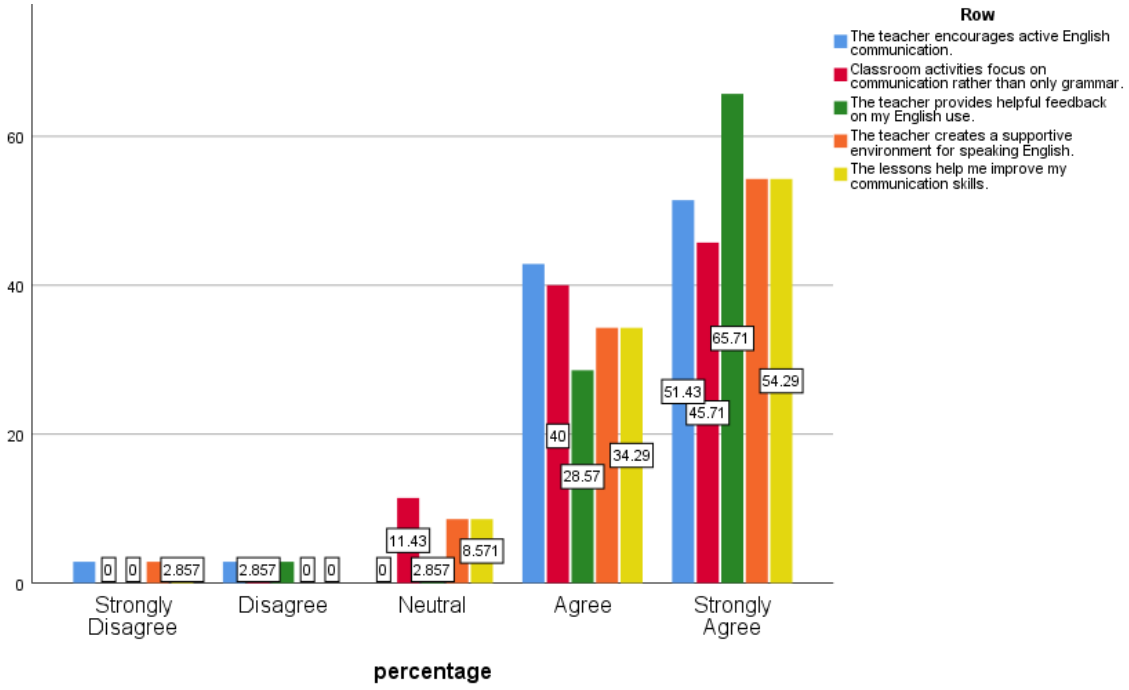
There are also positive results when it comes to the use of English by students during online lessons. Many students also stated that English was the primary language spoken in the classroom with 28.57 percent of students agreeing with this and 48.57 percent strongly agreeing (M=4.17, S.D.=1.014). In the same way, 17.14% answered in the affirmative and 34.29%

answered strongly in the affirmative that they more frequently used English than their first language in classes (M=3.71, S.D.=1.126). Concerning speaking opportunities, 45.71% had agreed and 40% strongly agreed that they had numerous opportunities to speak English online classes (M=4.2, S.D.=0.868). Besides, 31.43% of them agreed and 37.14% strongly agreed that they attempted to demonstrate their ideas in English when doing class tasks (M=3.94, S.D.=1.056). Lastly 37.14 percent said yes, and 40 percent said strongly that they were encouraged to talk English even when they made a mistake (M=4.09, S.D.=0.981). All in all, the percentages of agreement are high, which implies that the online learning environment facilitated active use of English among students in classes.

4.1.3 Teaching Practices and Communication

Figure 3.

Teaching Practices and Communication

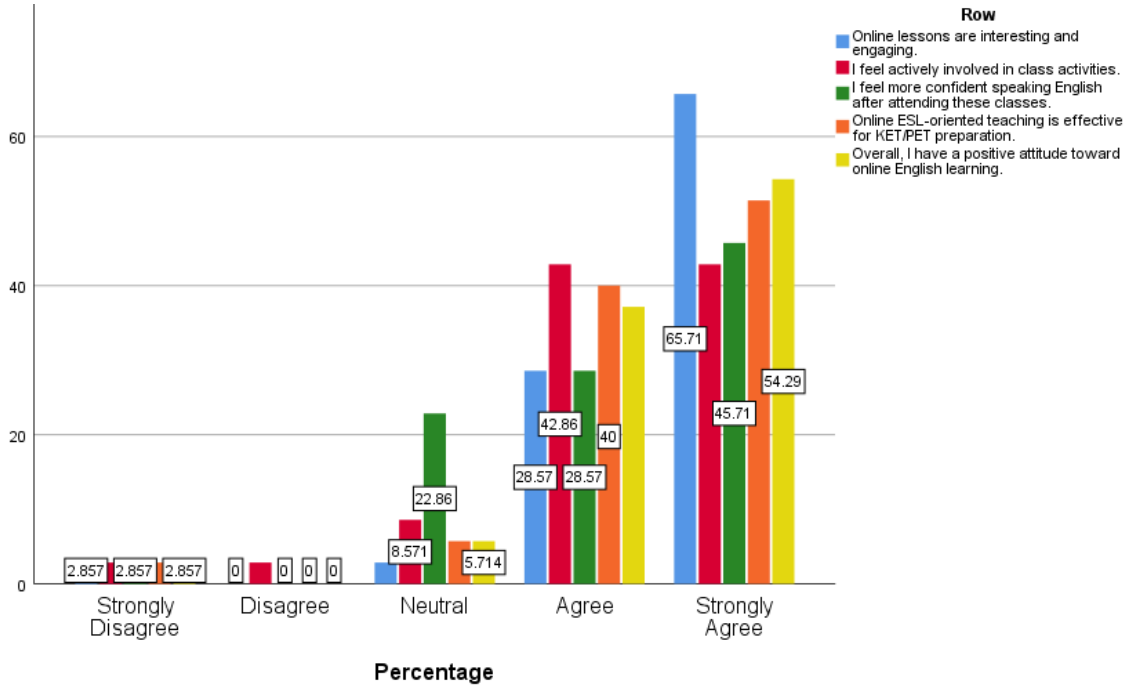


The findings also reveal favorable attitudes of instructional activities in the online English lesson. Most of the students indicated that the teacher supported active communication in English as 42.86% and 51.43% students were of the view that it was true and strongly true respectively (M=4.37, S.D.=0.877). On the same note, 40 percent of them said that they agree, and 45.71 percent strongly agreed that classroom practices centered on communication as opposed to grammar (M=4.29, S.D.=0.789). As far as feedback is concerned, 28.57% of the respondents had to agree and 65.71% strongly agreed that the teacher gave them useful feedback on how they used English (M=4.57, S.D.=0.698). Moreover, 34.29% said that the teacher provided a favorable environment to speak English and 54.29% said strongly (M=4.37, S.D.=0.877). Lastly, 34.29% concurred with the statement that the lessons assisted them in improving their communication skills and 54.29% strongly concurred with the statement (M=4.37, S.D.=0.877). In general, these results indicate that the instructional method was effective in helping students in the online classroom to develop in the communicative language.

4.1.4 Engagement and Overall Perception

Figure 4.

Engagement and Overall Perception



The findings also indicate good attitudes to online learning of English among the students as well as high engagement and effectiveness. Most students said that they found online lessons interesting and engaging with 28.57% agreeing and 65.71% strongly agreeing (M=4.54, S.D.=0.817). Likewise, 42.86% of them agreed and 45.71% strongly agreed that they were actively engaged in the activities in classes (M=4.20, S.D.=0.933). Regarding the aspect of speaking confidence, 28.57 percent answered that they agreed, and 45.71 percent strongly agreed that they felt much more confident speaking English after taking these classes (M=4.14, S.D.=0.974). Besides, 40 percent and 51.43 percent consented and strongly consented to the ESL-oriented teaching approach as effective in the preparation of KET/PET (M=4.37, S.D.=0.843). Lastly, 37.14% chose yes and 54.29% chose strongly yes that they had an overall positive attitude toward online learning of English (M=4.40, S.D.=0.847). Generally, the large percentages of agreement indicate that the online classes were effective in fostering engagement of the students, their confidence, and positive attitudes towards learning.

4.2 Findings from interviews

The thematic analysis was the method used to analyze qualitative data gained during semi-structured interviews. The responses of the students elicited two key themes: positive attitudes towards the online learning of English and supportive functions of digital tools in the online interaction and communication.

4.2.1 Positive attitudes towards the online learning of English

Majority of students indicated that they felt good about learning English via Google Meet. They added that online education was convenient as they could learn at home and at the same time communicate with their instructors and peers. Other students also said that they were comfortable attending online classes and working on speaking exercises. The opinions of some students are as follows:

“I think learning English on Google Meet is very convenient. I can study at home and still talk with my teacher and classmates.” (S1)

“I like learning English on Google Meet because the teacher explains clearly and I can see the slides easily.” (S2)

“For me, learning online is interesting. We can practice speaking and answering questions in class. (S3)

“I think Google Meet helps me feel comfortable when learning English. I am not afraid to speak in class.” (S4)

“I enjoy my online English class because the teacher gives many activities and we can practice English together.” (S5)

4.2.2 Supportive functions of digital tools in the online interaction and communication

Digital tools and the interaction and understanding process. Another aspect is that students emphasized the significance of digital tools applied in online lessons. The use of screen sharing, slides, and the chat box feature enabled the students to have a better comprehension of the lesson and engage in classroom activities. Some of the students said that these tools made the lessons more interesting and made them feel more comfortable speaking English. Indicatively, a student said, “The teacher shares slides on the screen, so I can understand the lesson better” (S1). “When the teacher shows pictures and exercises on the screen, it helps me practice English more easily” (S3). “The chat box helps me answer questions and ask the teacher when I do not understand.” (S2). Likewise, another student observed that Digital tools make it more entertaining, and I feel more engaged in the classroom “Digital tools make the lesson more interesting, and I feel more active in class” (S4). Lastly, there was also more confidence in communication “I think the tools help me speak and answer questions in English more confidently.” (S5).

In general, the interview findings indicated that students gave positive ratings on the perception of online English learning via Google Meet and thought that the digital tools were significant in facilitating interaction, engagement, and communication in online ESL-oriented courses.

5. Discussion

This study was aimed at exploring how ESL-oriented online English teaching by using Google Meet is perceived by learners in private classes on Cambridge exam preparation in Ho Chi Minh City. Moreover, the research examined the perceptions of students towards the use of digital tools on Google Meet to facilitate interaction, engagement, and communication during online classes.

RQ1: How do students perceive ESL-oriented online English learning through Google meet in private classes of Cambridge exam preparation in Ho Chi Minh City?

Concerning the general perceptions of online ESL-based learning among the students, the results showed that most of the students had positive perceptions about learning English using Google meet, with 45.71% feeling comfortable using the platform. They actively engaged with teachers and peers and positively viewed the digital tools' role in enhancing comprehension. Approximately 48.57% reported that English was used frequently in class, and many felt encouraged to communicate despite errors. Students expressed favorable attitudes towards instructional activities, appreciating teacher support and feedback effectiveness. Engagement levels were high, with 65.71% finding online lessons engaging and 45.71%

feeling more confident in their English-speaking abilities after classes. These findings were further supported by the interview data, in which students reported that learning through Google Meet was convenient and allowed them to communicate effectively. These findings are in line with the earlier studies by Yumnam (2021) and Purwoko et al. (2022) that have indicated that online learning environment can be used successfully as a means of facilitating communication and interaction in language learning settings.

RQ2: What is the perception of learners regarding digital tools utilization on Google Meet as an assistant to the processes of interaction, engagement, and communication in English concerning online ESL-oriented courses?

The results indicated that digital tools are very crucial in improving the learning experience of students. According to the questionnaire results, most students (around 82 percent) found the use of such tools as screen sharing, slides, and chat functions to be helpful in enhancing their understanding and engagement. This was supported by the answers of the interviews where the students said that these tools made lessons engaging, allowed them to follow the material easier and became more confident with their use of English. The results are consistent with the existing studies such as the study by Solihah (2023), suggesting that online learning can be supported using digital tools that can help to improve the engagement of learners, their interaction, and communication.

In general, the findings of the current paper indicate that a successful online learning experience can be achieved through the combination of communicative teaching methods and digital tools when using such a platform as Google Meet. This kind of environment may motivate students to use English actively, communicate with their teacher and colleagues, and attend online ESL-oriented classes with greater confidence.

6. Conclusion and Implications

The research examined the perceptions of students regarding ESL-based online teaching of English via Google Meet during the private classes in preparation of Cambridge exams in Ho Chi Minh City. The findings indicated that most of the students were positive about the online learning environment. Students said that they could communicate with their teacher and classmates, speak English in the classroom and feel encouraged to express their ideas. Moreover, online technologies like screensharing, slides, and chat applications assisted students in learning lessons more effectively and provided more interaction in a classroom. All in all, the results indicate that the application of Google Meet along with communicative teaching and digital tools could be effective as an effort to support the process of interaction, engagement, and communication in online ESL classes.

First, the sample size was not very large, consisting of 35 students of the private classes to prepare to Cambridge exams in Ho Chi Minh City, which might affect the generalization of the results. Second, the study involved questionnaires and interviews, but there were only a few participants in interviews, which means that the qualitative data might not be comprehensive and reflect the results of all students. Lastly, the research could not include any learners other than KET and PET; hence, the results are not necessarily applicable to learners with different proficiency levels.

On pedagogical level, educators are advised to develop communicative lessons and competently use digital technology to improve interaction and engage students in online learning. The following studies might include larger samples, learners of varying proficiency, and other types of data collection aimed at further investigation of the influence of online technologies on language learning of students.

References

Al-Jarf, R. (2026). A Self-Systematic Review of Mobile Apps for Developing Multiple Language Skills in EFL. *Journal of Computer Science and Technology Studies*, *8*(3), 14-29. https://doi.org/10.32996/jcsts.2026.8.3.2

Al-Maroof, R. S., Alshurideh, M. T., Salloum, S. A., AlHamad, A. Q. M., & Gaber, T. (2021). Acceptance of google meet during the spread of coronavirus by arab university students. Informatics, 8(2), 24. https://doi.org/10.3390/informatics8020024

Al-Maroof, R. S., Salloum, S. A., Hassanien, A. E., & Shaalan, K. (2020). Fear from COVID-19 and technology adoption: the impact of google meet during coronavirus pandemic. Interactive Learning Environments, 31(3), 1293-1308. https://doi.org/10.1080/10494820.2020.1830121

AlTameemy, F. A., & Alrefaee, Y. (2021). Impact of covid-19 on English language teaching in Yemen: Challenges and opportunities. *TESOL International Journal, 16*, 238-252.

Delviyanti, M., Ismiyati, Y., & Dinata, R. P. (2022). Students' perception of English studies uses the Google Meet application at the tenth grade of Senior High School Number 3 Jambi City. *JELT: Journal of English Language Teaching*, *6*(1), 73-79.

El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*, *18*(1), 53. https://doi.org/10.1186/s41239-021-00289-4

El-Seoud, M. S. A., Taj-Eddin, I. A., Seddiek, N., El-Khouly, M. M., & Nosseir, A. (2014). E-learning and students' motivation: a research study on the effect of e-learning on higher education. *International Journal of Emerging Technologies in Learning (iJET)*, 9(4), 20. https://doi.org/10.3991/ijet.v9i4.3465

Eraković, B., & Topalov, J. (2021). Teaching and learning through moodle, google doc and zoom: fostering student engagement in (a)synchronous learning environments. Inovacije u nastavi, 34(4), 122-136. https://doi.org/10.5937/inovacije2104122e

Ha, V. X., Tran, N. G., & Tran, N. H. (2021). Teachers' beliefs and practices regarding assessment in English as a Foreign Language classrooms in Vietnam. *The Qualitative Report*, *26*(11), 3457-3475. Article 8. https://doi.org/10.46743/2160-3715/2021.5063

Hsu, W. (2015). Transitioning to a communication-oriented pedagogy: Taiwanese university freshmen's views on class participation. *System*, *49*, 61-72. https://doi.org/10.1016/j.system.2014.12.002

Moorhouse, B. L., & Wong, K. M. (2021). Blending asynchronous and synchronous digital technologies and instructional approaches to facilitate remote learning. *Journal of Computers in Education*, 9(1), 51-70. https://doi.org/10.1007/s40692-021-00195-8

Nehe, B. M. (2021). Students' perception on Google Meet video conferencing platform during English speaking class in pandemic era. *English Review: Journal of English Education*, 10(1), pp. 93-104. https://doi: https://doi.org/10.25134/erjee.v10i1.5359

Pham, D. T., & Nguyen, T. H. H. (2023). University students' perceptions of google tools in learning english courses online. International Journal of Emerging Technologies in Learning (iJET), 18(13), 45-61. https://doi.org/10.3991/ijet.v18i13.39857

Popovici, A., & Mironov, C. (2015). Students' perception on using eLearning technologies. *Procedia-social and behavioral sciences*, *180*, 1514-1519. https://doi.org/10.1016/j.sbspro.2015.02.300

Procel, G. J. O., Medina, M. L. F., Sánchez, D. J. S., & Tacuri, M. A. P. (2024). Using technology in english teaching. CID - Centro de Investigación y Desarrollo. https://doi.org/10.37811/cli_w1048

Purwoko, R. Y., Primartadi, A., Efendi, Y., & Amin, T. (2022). The online learning creativity using the Google Meet platform in vocational high schools. *Tarbawi: Jurnal Ilmu Pendidikan*, *18*(2), 111-121. https://doi.org/10.32939/tarbawi.v18i1.1313

Radovan, A., Mihaljević, B., & Žagar, M. (2020). Experiences related to using online communication tools for distance learning. EDULEARN Proceedings, 1, 8386-8389. https://doi.org/10.21125/edulearn.2020.2063

Religioni, A., Meilinda, C., Alawiyah, A., & Farrel, L. M. (2024). Differentiated instructions in ESL and EFL classrooms: A systematic literature review. *STAIRS: English Language Education Journal*, *5*(2), 123-132.

Rintaningrum, R. (2023). Technology integration in English language teaching and learning: Benefits and challenges. *Cogent Education*, *10*(1). https://doi.org/10.1080/2331186X.2022.2164690

Sato, M., & Storch, N. (2020). Context matters: learner beliefs and interactional behaviors in an EFL vs. ESL context. *Language Teaching Research*, 26(5), 919-942. https://doi.org/10.1177/1362168820923582

Solihah, Y. (2023). Students' perception on the online learning through Google Meet and Youtube channel during pandemic situation. *SALEE: Study of Applied Linguistics and English Education, 4*(1), 171-184. https://doi.org/10.35961/salee.v4i1.609

Yumnam, R. (2021). E-learning: an effective mode of teaching English as a second language. *Journal of Translation and Language Studies*, *2*(2), 1–9. https://doi.org/10.48185/jtls.v2i2.275

Zahara, M. A., & Jupri, A. R. (2022). The effect of Microsoft PowerPoint interactive learning media application on reading comprehension skills of 3rd grade elementary school students for Indonesian language subjects. *Jurnal Paedagogy*, *9*(3), 532-539. https://doi.org/10.33394/jp.v9i3.5383

Zhang, B. (2023). The relationship between teacher-student rapport and EFL learners' engagement in online scaffolding setting. *Journal of Psycholinguistic Research*, 52(5), 1685-1705. https://doi.org/10.1007/s10936-023-09954-3

Zhou, J., Lee, S. S., Kew, S. N., & Liew, C. Y. (2024). Multiliteracies-translanguaging conceptual model framework for student agency in english as foreign language contexts. *International Journal of Evaluation and Research in Education (IJERE)*, 13(6), 4046. https://doi.org/10.11591/ijere.v13i6.29468

Bionote

Ngo Thi Thu Hien is a master's graduate in English Language from HUFLIT - Ho Chi Minh City University of Foreign Languages and Information Technology. She is now teaching online from home, including the Cambridge English program. She helps students build language competence and prepare for exams. She has also delivered a conference paper,

advancing research in English language teaching methods. Her research includes language teaching, motivation, and classroom techniques.

Appendix A

Questionnaire (English Version)

Learners' Perceptions of ESL-Oriented Online English Teaching via Google Meet: A Questionnaire-Based Study in Private Cambridge Exam Preparation Classes in Ho Chi Minh City

Hi everyone!

I am researching the Learners' Perceptions of ESL-Oriented Online English Teaching via Google Meet in Private Cambridge Exam Preparation Classes in Ho Chi Minh City. The findings can offer some information about the potential of using digital tools on Google Meet to help students feel confident and communicate within online ESL learning spaces, so I appreciate your collaboration. Thank you so much!

Part 1: Demographics

Please answer the following question for your personal information

1. Gender:

Male

Female

N/A

2. Age:

6-8

9-10

Over 10

3. How many years have you been learning English?

Less than 3 years

3-4 years

4-5 years

Over 5 years

4. How long have you been learning English online??

1 year

2-3 years

More than 3 years

Part 2: Learners' Perceptions of ESL-Oriented Online English Teaching via Google Meet

Please choose the appropriate number

1= Strong Disagree

2= Disagree

3= Neutral

4= Agree

5= Strong agree

	1	2	3	4	5
Google Meet & Digital Tools					
I feel comfortable learning English on Google Meet.					
I can interact easily with my teacher during online lessons.					
I can communicate effectively with classmates during online lessons.					
Digital tools (screen sharing, slides, chat box, etc.) help me understand the lesson better.					
The use of digital tools makes lessons more interactive.					
	1	2	3	4	5
ESL-Oriented Language Use					
English is the main language used during online lessons.					
I use English more often than my first language in class.					
I have many opportunities to speak English during lessons.					
I try to express my ideas in English during class activities.					
I feel encouraged to speak English even when I make mistakes.					
Teaching Practices and Communication	1	2	3	4	5
The teacher encourages active English communication.					
Classroom activities focus on communication rather than only grammar.					
The teacher provides helpful feedback on my English use.					
The teacher creates a supportive environment for speaking English.					
The lessons help me improve my communication skills.					

Engagement and Overall Perception	1	2	3	4	5
Online lessons are interesting and engaging.					
I feel actively involved in class activities.					
I feel more confident speaking English after attending these classes.					
Online ESL-oriented teaching is effective for KET/PET preparation.					
Overall, I have a positive attitude toward online English learning.					

Thank you for your cooperation!

APPENDIX B

QUESTIONNAIRE (Vietnamese version)

Đề tài: Nhận thức của người học về việc dạy tiếng Anh trực tuyến theo định hướng ESL qua Google Meet: Một nghiên cứu dựa trên bảng hỏi trong các lớp chuẩn bị kỳ thi Cambridge tư nhân tại Thành phố Hồ Chí Minh

Chào các bạn học viên!

Cô đang nghiên cứu về nhận thức của người học về giảng dạy Tiếng Anh trực tuyến theo định hướng ESL qua Google Meet trong các lớp chuẩn bị kỳ thi Cambridge tư nhân tại Thành phố Hồ Chí Minh. Kết quả của nghiên cứu có thể cung cấp một số thông tin về tiềm năng của việc sử dụng các công cụ kỹ thuật số trên Google Meet để giúp học sinh cảm thấy tự tin và giao tiếp trong các không gian học tập ESL trực tuyến, vì vậy cô rất trân trọng sự hợp tác của bạn. Cảm ơn các bạn rất nhiều!

Phần 1: Thông tin cá nhân

Vui lòng trả lời câu hỏi sau để cung cấp thông tin cá nhân của bạn.
1. Giới tính:
Nam
Nữ
Không xác định
2. Độ tuổi:
6-8
9-10
Trên 10
3. Bạn đã học tiếng Anh được bao nhiêu năm?
Ít hơn 3 năm
3-4 năm
4-5 năm
Hơn 5 năm
4. Bạn đã học tiếng Anh trực tuyến được bao lâu rồi?
1 năm
2-3 năm
Hơn 3 năm
Phần 2: Nhận thức của người học về việc dạy tiếng Anh trực tuyến theo định hướng ESL qua Google Meet
Vui lòng chọn số thích hợp
1= Rất không đồng ý
2= Không đồng ý
3= Trung lập
4= Đồng ý
5= Rất đồng ý

	1	2	3	4	5
Google Meet & Công Cụ Kỹ Thuật Số					
Tôi cảm thấy thoải mái khi học tiếng Anh trên Google Meet.					

Tôi có thể tương tác dễ dàng với giáo viên của mình trong các bài học trực tuyến.					
Tôi có thể giao tiếp hiệu quả với các bạn học trong các bài học trực tuyến.					
Các công cụ kỹ thuật số (chia sẻ màn hình, slide, hộp chat, v.v.) giúp tôi hiểu bài học tốt hơn.					
Công cụ số giúp giờ học thú vị hơn.					
Sử Dụng Ngôn Ngữ Hướng Đến ESL	1	2	3	4	5
Tiếng Anh là ngôn ngữ chính được sử dụng trong các bài học trực tuyến.					
Tôi sử dụng tiếng Anh thường xuyên hơn so với ngôn ngữ mẹ đẻ của mình trong lớp học.					
Tôi có nhiều cơ hội để nói tiếng Anh trong các tiết học.					
Tôi cố gắng diễn đạt ý tưởng của mình bằng tiếng Anh trong các hoạt động trên lớp.					
Tôi cảm thấy tự tin nói tiếng Anh dù có sai.					
Thực hành giảng dạy & Giao tiếp	1	2	3	4	5
Cô khuyến khích tôi nói tiếng Anh nhiều.					
Hoạt động trên lớp chủ yếu để giao tiếp, không chỉ học ngữ pháp.					
Cô sửa và góp ý giúp tôi nói tiếng Anh tốt hơn.					
Cô tạo không khí thoải mái để tôi nói tiếng Anh.					
Bài học giúp tôi giao tiếp tiếng Anh tốt hơn.					
Sự tham gia & Nhận thức tổng thể	1	2	3	4	5
Các bài học trực tuyến thú vị và hấp dẫn.					
Tôi cảm thấy mình tham gia tích cực vào các hoạt động trên lớp.					

Tôi cảm thấy tự tin hơn khi nói tiếng Anh sau khi tham gia các lớp học này.					
Giáo dục trực tuyến hướng đến ESL hiệu quả cho việc chuẩn bị thi KET/PET.					
Nhìn chung, tôi có thái độ tích cực đối với việc học tiếng Anh trực tuyến.					

Cảm ơn sự hợp tác của các bạn!

Appendix C

Interview Questions

1. How do you feel about learning English via Google Meet in your Cambridge exam preparation course?
2. What do you think of digital tools in Google Meet (screen sharing, slides, chat) that enable you to interact, participate, and communicate in English in an online lesson?

#42 Examining the Relationship between AI-Supported Self-Regulated Learning and Speaking Self-Efficacy among Vietnamese EFL University Students

Hoang The Hai Nguyen

Abstract

Artificial intelligence (AI) has increasingly become integrated into English as a foreign language (EFL) education, providing learners with new chances to plan, monitor, and regulate their language learning. However, the relationship between AI-supported self-regulated learning (AI-SRL) in speaking skills and learners' speaking self-efficacy remains underexplored, particularly in Vietnamese higher education. This study adopted a cross-sectional quantitative correlational design to examine the reported levels of AI-SRL and speaking self-efficacy and the relationship between the two constructs among 186 EFL students at a private university in Vietnam. A structured questionnaire was used for data collection and analyzed by descriptive statistics, reliability analysis, exploratory factor analysis and pearson's product–moment correlation. The results showed that both AI-SRL and speaking self-efficacy were reported at levels above the midpoint of the five-point scale. A statistically significant moderate positive correlation was also found between the two constructs ($r = .458$, $p < .001$), indicating that students who reported stronger regulation of their AI-supported speaking practice also tended to report stronger beliefs in their capability to perform English-speaking tasks. Because the study employed a cross-sectional correlational design, the findings do not establish a causal relationship between AI-SRL and speaking self-efficacy. Instead, they provide context-specific evidence of an association between learners' reported regulatory behaviour during AI-supported speaking practice and their perceived speaking capability.

Keywords: AI-supported self-regulated learning, speaking self-efficacy, EFL university students, artificial intelligence, Vietnamese higher education

1. Introduction

Artificial intelligence (AI) has increasingly become visible in English as a foreign language (EFL), providing learners with new chances for interaction, feedback, and language practice (Kohnke et al., 2023; Lai & Lee, 2024). Speaking is one of the most useful aspects. Recent studies of AI-supported speaking have reported positive outcomes related to speaking skills, willingness to communicate, confidence, and learner engagement (Du & Daniel, 2024; Fathi et al., 2024). Generative AI can literally support interactive exchanges and provide language-related feedback, while automated speech recognition (ASR) can give learners information based on their spoken production (Kohnke et al., 2023; Sun, 2023). For example, Sun (2023) found that ASR-supported practice was associated with improvements in pronunciation and broader speaking skills among EFL learners. Together, these findings suggest that AI can extend opportunities for learners to practise and review spoken English beyond regular classroom interaction.

AI-supported speaking practice may also shape how learners experience oral communication (Fathi et al., 2024; Wang et al., 2024; Zhang et al., 2024). Zhang et al. (2024) found that EFL learners practicing with an AI-speaking assistant showed lower foreign language anxiety and greater willingness to communicate after the intervention. Similar findings have been reported in research on conversational generative AI, although outcomes

may vary depending on the type of tool and interaction provided (Wang et al., 2024). These studies suggest that AI-supported environments can provide useful conditions for repeated speaking practice.

Chang and Sun's (2024) systematic review showed that AI has been used to support different stages of self-regulated language learning, while more recent work has examined how EFL learners plan, use strategies, monitor progress, and reflect in AI-supported environments (Hou & Zhou, 2025). These findings point to an important distinction between simply using an AI tool and regulating learning while using it. For example, one learner may use AI mainly to obtain an answer or correction, whereas another may set a goal, examine feedback, compare attempts, and decide what to practise next. The difference lies partly in how learners manage the learning process.

Self-regulated learning (SRL) provides a useful framework for examining this difference. Recent research describes SRL as a proactive and adaptive process in which learners plan and adjust their thoughts and actions in relation to learning goals (Pan et al., 2025). In AI-supported language learning, regulation may involve goal setting, strategic use of AI resources, monitoring, feedback evaluation, and reflection across different stages of learning (Chang & Sun, 2024; Hou & Zhou, 2025). These processes are cyclical rather than limited to a single learning event, as information gained from one attempt can influence how learners approach later practice (Chang & Sun, 2024).

Within the speaking context, these self-regulatory processes are directly translated into strategic learning decisions (Zimmerman, 2000, 2002). An autonomous learner may diagnose specific oral difficulties, select appropriate conversational AI tasks, monitor online speaking fluency, evaluate automated feedback, and refine speaking strategies for future practice behaviors that are increasingly prominent in AI-mediated EFL speaking environments (Du & Daniel, 2024; Panadero, 2017; Zimmerman, 2002). Although conversational AI provides rich opportunities for oral communication, the ultimate learning gains depend on how students interpret and utilize technological feedback, highlighting the need for active strategic control over digital resources (Hou & Zhou, 2025; Lai & Gu, 2011; Tseng et al., 2006). Therefore, this study defines AI-supported self-regulated learning (AI-SRL) as the extent to which learners plan, monitor, evaluate, and adjust their AI-supported English-speaking practice, adapting classic SRL frameworks to contemporary generative AI learning contexts (Chang & Sun, 2024; Pintrich, 2000; Su et al., 2023)."Speaking self-efficacy is another learner-related factor relevant to this process. In second-language research, self-efficacy refers to learners' beliefs about their capability to perform particular language tasks successfully (Li & Zhang, 2023; Phipps, 2024). It is therefore different from objectively measured language ability. Learners with similar levels of speaking performance may not necessarily evaluate their own speaking capability in the same way. Phipps (2024), for example, found that speaking self-efficacy among university EFL learners changed over time and was related to how learners evaluated speaking outcomes. Li and Zhang (2023) also reported dynamic relationships between speaking self-efficacy and the development of speaking accuracy and fluency in online-supported learning. These findings indicate that speaking self-efficacy can develop through learners' experiences of speaking practice and their interpretations of those experiences.

There is also a theoretical reason to consider speaking self-efficacy together with self-regulation. Derakhshan and Fathi (2024) examined self-efficacy, self-regulation, and L2 speaking performance within the same model and found that both learner factors were meaningfully connected with speaking performance. Although their study did not focus on AI-supported speaking, it suggests that self-regulation and self-efficacy can be examined together

within the broader process of L2 speaking development. Research on AI-supported SRL adds another dimension by showing that learners may plan, monitor, and reflect while using AI-supported resources for language learning (Chang & Sun, 2024; Hou & Zhou, 2025).

A possible connection between AI-SRL and speaking self-efficacy may emerge through repeated speaking practice. Monitoring and feedback evaluation can provide learners with information about what has improved and what remains difficult. Such information may become relevant when learners judge what they believe they are capable of doing in later speaking tasks. At the same time, learners' existing speaking self-efficacy may also be related to how they approach practice, respond to difficulty, and evaluate their learning experiences (Li & Zhang, 2023; Phipps, 2024). The relationship may therefore operate in more than one direction. For this reason, the present study does not assume that stronger AI-SRL leads to higher speaking self-efficacy or that higher speaking self-efficacy leads to stronger AI-SRL. Instead, it examines whether the two constructs are statistically associated.

Existing research has addressed different parts of this issue, but the specific relationship remains underexplored. Studies of conversational AI and AI-supported speaking have mainly focused on speaking outcomes, learner perceptions, anxiety, willingness to communicate, and the affordances of AI tools (Du & Daniel, 2024; Fathi et al., 2024; Lai & Lee, 2024; Zhang et al., 2024). Another line of research has begun to examine self-regulation in AI-supported language learning (Chang & Sun, 2024; Hou & Zhou, 2025; Pan et al., 2025). Speaking research has separately examined self-efficacy and its relationship with speaking development, performance, and self-regulation (Derakhshan & Fathi, 2024; Li & Zhang, 2023; Phipps, 2024). Although these studies provide important foundations for the present research, it remains less clear whether learners who report stronger self-regulation specifically during AI-supported speaking practice also report stronger speaking self-efficacy.

This question is particularly relevant in the Vietnamese EFL context. Duong and Suppasetseree (2024) investigated AI voice-chatbot practice among Vietnamese undergraduate students and reported significant improvement in English-speaking skills after an eight-week intervention. Their study demonstrates that AI-supported speaking has already become a relevant area of research in Vietnamese higher education. Ngo et al. (2025) also examined Vietnamese EFL undergraduates' AI use, self-efficacy, and self-directed learning and reported a positive relationship between self-efficacy in using AI and self-directed learning with AI.

However, the constructs examined in these studies differ from those used in the present research. Duong and Suppasetseree (2024) focused on speaking performance following AI-supported practice rather than learners' self-regulation during that practice. Cong-Lem et al. (2025) examined Vietnamese EFL learners’ self-efficacy in using generative AI tools, whereas the present study focuses specifically on learners’ perceived capability to perform English-speaking tasks with AI support.Their self-directed learning construct also addressed AI-supported learning more broadly, while AI-SRL in the present study is limited specifically to the regulation of AI-supported English-speaking practice. Being confident in using an AI tool is not necessarily the same as believing that one can perform a speaking task successfully. Likewise, general independent use of AI does not necessarily show how a learner plans, monitors, evaluates, and adjusts speaking practice.

Within the recent literature reviewed for the present study, direct quantitative evidence on the relationship between AI-supported self-regulated learning in speaking practice and speaking self-efficacy remains limited, particularly in Vietnamese university EFL contexts (Cong-Lem et al., 2025; Duong & Suppasetseree, 2024). The present study therefore focuses on this specific relationship rather than asking only whether students use AI or whether AI-

supported practice improves speaking performance, maintaining the critical distinction between technology usage and active self-regulatory control (Chang & Sun, 2024; Su et al., 2023). It examines how students report regulating their AI-supported speaking practice and whether this overall regulatory behaviour is associated with their beliefs about their capability to perform English-speaking tasks (Derakhshan & Fathi, 2024; Zimmerman, 2000).

The study was conducted with EFL students at a private university in Vietnam. It first examined students' reported levels of AI-SRL and speaking self-efficacy and then investigated the statistical relationship between the two constructs. Because the study employed a cross-sectional correlational design, it did not attempt to establish that one variable causes changes in the other (Creswell & Creswell, 2018; Slater & Hasson, 2025). The focus was limited to whether AI-SRL and speaking self-efficacy were associated and, if so, the direction and strength of that association (Field, 2024).

Accordingly, the study addressed the following research questions:

What are the reported levels of AI-supported self-regulated learning and speaking self-efficacy among EFL students at a private university in Vietnam?

Is there a statistically significant relationship between AI-supported self-regulated learning and speaking self-efficacy among these students?

2. Literature Review

2.1. Artificial Intelligence in Language Learning

Artificial intelligence has provided EFL learners with more opportunities for interaction, feedback, language production, and independent practice (Godwin-Jones, 2023; Kohnke et al., 2023; Lo, 2023). For speaking practice, two forms of technology are particularly relevant to the present study: generative AI and automated speech recognition (ASR). Generative AI can support interactive language practice by responding to learners' input, while ASR processes spoken language and provides information about learners' oral production (Kohnke et al., 2023; Sun, 2023). Both technologies can extend speaking practice beyond scheduled classroom interaction.

Recent studies have shown that AI-supported activities can be useful for speaking development. Fathi et al. (2024), for example, reported positive effects of AI-mediated interaction on EFL learners' speaking skills and willingness to communicate. ASR-supported practice has also been associated with improvements in pronunciation and speaking skills (Sun, 2023). These findings suggest that AI can provide useful conditions for repeated speaking practice. The availability of such opportunities, though, does not explain how learners actually use them.

Generative AI can support several parts of language practice, such as generating ideas, suggesting language, providing feedback, and creating conversational exchanges (Kohnke et al., 2023; Lo, 2023). Learners can also adjust prompts according to what they want to practise, which gives them some control over the direction of an activity. This flexibility may be useful when preparing for presentations, interviews, discussions, or other speaking situations. AI-supported practice is therefore less fixed than many traditional materials, but its learning value may still depend on the learner's level of engagement.

The availability of AI does not necessarily mean that learners regulate their learning actively (Chang & Sun, 2024; Hou & Zhou, 2025). Recent research on AI-supported language learning shows that self-regulated engagement involves processes such as goal setting, strategic

use of learning resources, monitoring, and reflection (Chang & Sun, 2024; Hou & Zhou, 2025). These processes require learners to make decisions about how AI is used and how information from the tool is incorporated into later learning (Hou & Zhou, 2025). For the present study, this distinction is important because AI use and AI-supported self-regulation are treated as related but different aspects of learning.

ASR can provide learners with information about their spoken production through speech recognition and transcription (Sun, 2023). Sun (2023) found that ASR-supported practice, combined with peer correction, contributed to improvements in EFL learners' pronunciation and speaking skills. During the intervention, learners reviewed ASR transcriptions, identified possible problems in their spoken production, and used this information during later practice (Sun, 2023). This process gave learners opportunities to notice aspects of their speech that required further attention. For the present study, ASR is therefore relevant as a source of performance information that learners may use during self-regulated speaking practice.

AI-supported speaking may also influence learners' affective experience during oral communication. Zhang et al. (2024) found that an AI-speaking assistant was associated with increased willingness to communicate and reduced foreign language anxiety among Chinese EFL university students. Wang et al. (2024) reported a similar pattern for a conversational GenAI chatbot with human-like avatars, with significant improvement in willingness to communicate and self-perceived communicative competence, together with lower foreign language speaking anxiety. These findings suggest that some AI-mediated speaking environments may provide a more supportive setting for oral practice (Wang et al., 2024; Zhang et al., 2024). The effects, however, appear to depend on the form of AI-supported interaction rather than on AI use alone. Wang et al. (2024), for example, found different outcomes across chatbot conditions, and no significant differences in speaking performance were observed across the groups.

2.2. Self-Regulated Learning

Self-regulated learning (SRL) refers to the processes through which learners actively manage their learning in relation to particular goals (Panadero, 2017; Pintrich, 2000; Zimmerman, 2002). Zimmerman’s cyclical model, which provides the main SRL framework for the present study, conceptualises regulation through three interconnected phases: forethought, performance, and self-reflection (Zimmerman, 2000, 2002). Although these sources provide the foundational framework, the cyclical model continues to inform recent research on technology-enhanced and AI-supported language learning (Chang & Sun, 2024; Hou & Zhou, 2025; Pan et al., 2025).

SRL involves more than studying independently. Learners need to establish goals, select appropriate strategies, monitor their progress, evaluate outcomes, and adjust their approach when necessary (Pan et al., 2025; Zhang & Zou, 2024). Recent second-language research describes self-regulation as an active and adaptive process in which strategies are used in relation to particular learning goals and conditions rather than simply possessed by learners (Pan et al., 2025). The important issue is therefore not only whether learners know a particular strategy, but whether they can select and use it appropriately when responding to a specific learning task (Zhang & Zou, 2024).

Self-regulation may also vary according to the task and learning environment. Zhang and Zou (2024) emphasise the interaction among learners, strategies, tasks, and learning contexts in self-regulated second-language learning. Preparing a presentation, practising

pronunciation, and participating in spontaneous conversation, for example, may place different demands on planning, monitoring, strategy use, and feedback evaluation (Panadero, 2017; Zimmerman, 2002). Technology-enhanced SRL research similarly suggests that regulatory processes can vary according to the learning activity and the type of technological support available (Pan et al., 2025). SRL is therefore better understood as a context-sensitive learning process than as a fixed pattern of behaviour applied in the same way across all situations.

Within Zimmerman's model, the forethought phase occurs before or at the beginning of a learning activity (Zimmerman, 2000, 2002). Learners identify goals, consider task demands, and decide how they intend to approach the activity (Zimmerman, 2000, 2002). In speaking practice, this may involve identifying an aspect of speaking that needs attention and selecting a suitable activity or resource. The goal established at this stage provides a reference point for later monitoring and evaluation.

The performance phase concerns what learners do while carrying out the activity. This phase includes strategy use, attention, and monitoring of ongoing performance (Zimmerman, 2000, 2002). During speaking practice, learners may attend to pronunciation, fluency, lexical choice, or their ability to communicate an intended meaning. When they notice that a strategy or approach is not producing the intended result, they may need to make adjustments while the activity is still in progress.

The self-reflection phase follows the learner's experience of the task. Learners evaluate what occurred, interpret the outcome, and use this information when considering later learning activities (Zimmerman, 2000, 2002). Reflection gives the model its cyclical character because information from one learning attempt can influence planning and strategy selection in a later attempt. Recent research on AI-supported and technology-enhanced SRL continues to describe regulation in terms of connected processes involving planning, action, monitoring, evaluation, and later adjustment (Chang & Sun, 2024; Pan et al., 2025).

Technology can create additional opportunities for these processes while also increasing the number of decisions learners need to make. Pan et al. (2025) found that technology can function as a learning environment, a resource, or a means of supporting self-regulation in second-language learning. When AI is introduced into this process, learners may need to decide which tool or activity is appropriate for a particular goal, which feedback deserves attention, and how information produced by the technology should influence later practice. These decisions provide the basis for examining SRL specifically within AI-supported language learning.

2.3. AI-Supported Self-Regulated Learning

In the present study, AI-supported self-regulated learning (AI-SRL) refers to the application of self-regulatory processes when learners use AI resources to support their English-speaking practice. Recent research has described self-regulation in AI-supported language learning in terms of planning, strategy use, monitoring, reflection, and adjustment while learners interact with AI-supported resources (Chang & Sun, 2024; Hou & Zhou, 2025). AI-SRL is therefore not treated here as a separate learning theory. Rather, the term is used to describe how established SRL processes operate within AI-supported speaking practice.

Recent research provides support for examining learning in this way. Chang and Sun's (2024) systematic review showed that AI has been used to support different stages of self-regulated language learning and highlighted the role of strategic AI use within a cyclical learning process. Hou and Zhou (2025) similarly identified regulatory processes involving planning, strategy execution, monitoring, reflection, and technological adaptation among EFL

learners in AI-supported environments. Together, these studies suggest that understanding AI-supported learning requires attention not only to what technological tools can provide, but also to how learners make decisions while using them.

In AI-supported speaking practice, regulation can occur across different stages of learning. Before an activity, learners may establish a speaking goal and decide how an AI resource can support that goal. During practice, they may apply strategies, monitor their performance, and respond to information generated by the AI system. After the activity, they may evaluate the feedback and their own performance before deciding what should be addressed in a later attempt (Chang & Sun, 2024; Hou & Zhou, 2025). In this way, AI-supported practice can form part of a continuing regulatory cycle rather than a series of isolated interactions with technology.

The distinction between AI use and AI-SRL is therefore central to the present study. Two learners may use the same AI tool but manage their speaking practice differently. One learner may mainly request corrections or model answers, whereas another may establish a specific speaking goal, examine feedback, compare different attempts, and decide what to practise next. Frequency of AI use alone cannot indicate how actively a learner regulates the learning process. For this reason, AI-use background information and AI-SRL are treated separately in the present research.

Technology-enhanced language-learning research provides further support for this distinction. Pan et al. (2025) showed that technology can support SRL in several ways, including providing environments for practice, supporting strategy use, and creating opportunities for learners to monitor and regulate their learning. However, technological support does not represent a single form of self-regulation across all learning situations. Its role depends partly on how learners incorporate the available resources into their own regulatory processes.

Based on Zimmerman's cyclical framework and recent AI-supported language-learning research, AI-SRL in the present study is conceptually discussed in relation to processes such as planning, performance monitoring, feedback evaluation, reflection, and subsequent adjustment (Chang & Sun, 2024; Hou & Zhou, 2025). These processes provide a theoretical basis for the content of the AI-SRL questionnaire items. However, they are treated as conceptual aspects of AI-SRL rather than empirically confirmed dimensions in the present study. The primary descriptive and correlational analyses therefore examine AI-SRL as an overall construct rather than analysing individual SRL phases or regulatory processes separately.

2.4. Speaking Self-Efficacy

Speaking self-efficacy concerns learners' beliefs about their capability to perform particular English-speaking tasks (Bandura, 1997, 2006; Phipps, 2024). Self-efficacy refers to perceived capability rather than an objective measure of actual performance (Bandura, 1997, 2006). Recent L2 speaking research has also treated self-efficacy as learners' beliefs about their capability in specific aspects of speaking performance (Li & Zhang, 2023; Phipps, 2024). For the present study, this distinction is important because the questionnaire measures what students believe they can do in speaking situations rather than their objectively assessed speaking proficiency.

Perceived capability and actual speaking performance are therefore related but should not be treated as the same construct (Bandura, 1997, 2006; Li & Zhang, 2023). Learners may differ in how they judge their speaking capability even when their language performance is

relatively similar, since efficacy judgements are based on how individuals interpret their abilities and experiences (Bandura, 1997, 2006). Recent longitudinal evidence also shows that speaking-related self-efficacy can change over time. Li and Zhang (2023), for example, tracked speaking accuracy, fluency, and self-efficacy across six observations and found that self-efficacy for both accuracy and fluency developed over the course of online-supported learning. Phipps (2024) similarly examined changes in speaking self-efficacy among university EFL learners and its relationship with learners' causal attributions and satisfaction with speaking outcomes. These findings suggest that speaking self-efficacy is not simply a fixed judgement of ability; it can develop as learners gain and interpret further speaking experiences.

Self-efficacy is also task- and domain-specific rather than a general sense of confidence (Bandura, 1997, 2006). This specificity is reflected in recent speaking research. Li and Zhang (2023), for instance, measured self-efficacy for speaking accuracy and self-efficacy for speaking fluency separately, showing that efficacy beliefs can be examined in relation to particular aspects of oral performance. A learner may therefore judge their capability differently across speaking demands, such as producing accurate language, speaking fluently, delivering a prepared presentation, or responding spontaneously in a discussion (Bandura, 2006; Li & Zhang, 2023). In the present study, speaking self-efficacy is consequently treated as learners' perceived capability to perform the English-speaking tasks represented in the questionnaire rather than as general confidence in English.

Recent L2 speaking research also provides evidence for considering self-efficacy together with self-regulation. Derakhshan and Fathi (2024) examined growth mindset, self-efficacy, self-regulation, and L2 speaking performance among EFL learners. Their structural model showed a direct positive relationship between self-efficacy and L2 speaking performance, while self-regulation also played a mediating role in the model (Derakhshan & Fathi, 2024). The study did not examine AI-supported speaking practice, so its findings cannot be treated as direct evidence for the relationship investigated in the present study. Its relevance lies in showing that self-efficacy and self-regulation can operate within the same broader process of L2 speaking development.

Bandura's theoretical framework identifies four main sources from which efficacy beliefs may develop: mastery experiences, vicarious experiences, social persuasion, and physiological and affective states (Bandura, 1997, 2006). The influence of these sources depends partly on how individuals interpret the information arising from their experiences rather than on the experience alone (Bandura, 1997). These four sources are included here to explain the theoretical background of speaking self-efficacy, not as separate variables in the present study. The current study measures learners' overall beliefs about their speaking capability and does not examine the individual contribution of the four efficacy sources.

2.5. Linking AI-Supported Self-Regulated Learning and Speaking Self-Efficacy

AI-SRL and speaking self-efficacy represent conceptually different aspects of the learning process. AI-supported self-regulation concerns how learners plan, monitor, evaluate, and adjust their learning when AI resources are involved (Chang & Sun, 2024; Hou & Zhou, 2025), whereas speaking self-efficacy concerns learners' beliefs about their capability to perform particular speaking tasks (Li & Zhang, 2023; Phipps, 2024). Recent L2 speaking research has also examined self-efficacy and self-regulation within the same learning context, suggesting that the two constructs can be related while remaining theoretically distinct (Derakhshan & Fathi, 2024). For the present study, this distinction allows the relationship between learners' regulatory behaviour and their perceived speaking capability to be examined without treating them as the same construct.

Self-regulatory activities can provide learners with information about their performance and progress. Research on AI-supported SRL shows that monitoring and reflection allow learners to review learning processes, evaluate outcomes, and use information from earlier activity when making later decisions (Chang & Sun, 2024; Hou & Zhou, 2025). During AI-supported speaking practice, this may involve setting a goal, completing a speaking activity, examining feedback, comparing attempts, and deciding whether later practice needs to change (Chang & Sun, 2024; Hou & Zhou, 2025). Recent speaking research also shows that self-efficacy can develop over time alongside learners' speaking experiences and performance outcomes (Li & Zhang, 2023; Phipps, 2024). Information noticed during regulated practice may therefore become relevant when learners judge what they believe they can accomplish in later speaking tasks, although this possibility does not establish a causal effect of AI-SRL on speaking self-efficacy.

The connection may also be considered from the perspective of learners' existing efficacy beliefs. Recent L2 research has reported relationships among self-efficacy, self-regulation, and speaking performance, indicating that these learner factors can operate within the same broader speaking-learning process (Derakhshan & Fathi, 2024). Li and Zhang (2023) also found dynamic relationships between speaking self-efficacy and the development of speaking performance, while Phipps (2024) showed that speaking self-efficacy changed over time and was related to how learners interpreted speaking outcomes. These findings leave open the possibility that learners' existing beliefs about their speaking capability may be associated with how they approach and manage later practice. They do not, however, demonstrate that stronger self-efficacy causes stronger self-regulation.

The available literature therefore provides a theoretical and empirical basis for examining AI-SRL and speaking self-efficacy together, but it does not establish the direction of the relationship between them. The present study consequently treats their relationship as an association rather than a causal pathway. It does not propose that stronger AI-SRL causes stronger speaking self-efficacy or that higher speaking self-efficacy causes stronger AI-SRL. The empirical question is whether learners who report higher levels of AI-SRL also tend to report higher levels of speaking self-efficacy.

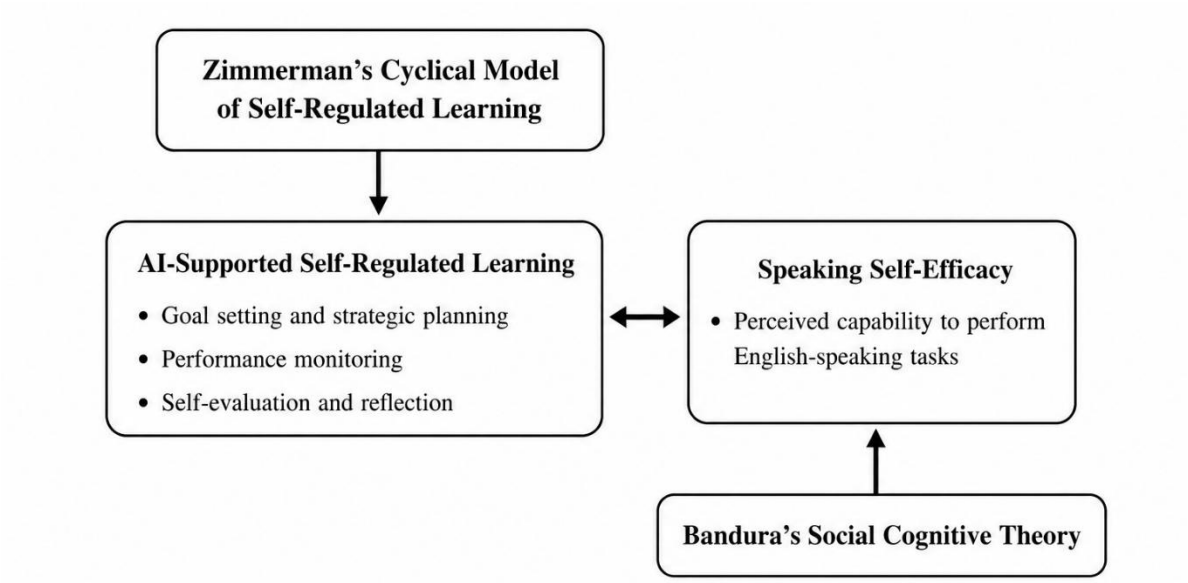
2.6. Theoretical Framework and Research Gap

The theoretical framework brings together Zimmerman's cyclical view of self-regulated learning and Bandura's concept of self-efficacy. Zimmerman's framework provides the theoretical basis for examining learners' planning, monitoring, evaluation, reflection, and adjustment of AI-supported speaking practice. Bandura's framework provides the basis for treating speaking self-efficacy as learners' perceived capability to perform speaking tasks. The two constructs remain separate, but both concern learner-related processes that may operate during speaking practice.

The relationship proposed in the present study is represented in Figure 1.

Figure 1

Theoretical Framework Linking AI-Supported Self-Regulated Learning and Speaking Self-Efficacy



Note. The double-headed arrow represents a theoretical and statistical association, not a causal pathway.

The double-headed arrow between AI-SRL and speaking self-efficacy represents the association examined in the present study rather than a causal relationship. The figure does not suggest that one construct produces changes in the other.

Studies of the use of AI in language learning have explored speaking performance, pronunciation, feedback, language anxiety, willingness to communicate, and learners' experiences of AI-mediated speaking practices (Fathi et al., 2024; Sun, 2023; Wang et al., 2024; Zhang et al., 2024). The study of language learning supported by AI and technology-enhanced language learning has also explored how learners plan, monitor, evaluate and reflect on their learning process with the presence of technological tools or resources (Chang & Sun, 2024; Hou & Zhou, 2025; Pan et al., 2025). A related line of L2 speaking research has considered self-efficacy in relation to speaking development, performance, and self-regulation (Derakhshan & Fathi, 2024; Li & Zhang, 2023; Phipps, 2024). These areas provide the main components of the present framework, but they have generally addressed different parts of the learning process rather than the specific relationship examined in this study.

This distinction becomes clearer when AI-supported speaking is considered. AI-mediated activities have been found to provide opportunities for speaking practice and to support aspects of speaking performance and willingness to communicate (Fathi et al., 2024; Sun, 2023). At the same time, research on AI-supported SRL shows that learners can engage in planning, strategy use, monitoring, and reflection while learning with AI (Chang & Sun, 2024; Hou & Zhou, 2025). Speaking research separately indicates that self-efficacy is related to learners' speaking development and performance (Derakhshan & Fathi, 2024; Li & Zhang, 2023; Phipps, 2024). What remains less clear is how these areas come together: specifically, whether learners who regulate their AI-supported speaking practice more strongly also tend to report stronger beliefs in their own speaking capability.

Recent research in Vietnam provides an important context for this question. Duong and Suppasetseree (2024) investigated an AI voice chatbot with Vietnamese undergraduate

students and found a significant improvement in English-speaking skills after an eight-week intervention. Their study focused primarily on speaking outcomes following AI-supported practice rather than on learners' self-regulation during that practice. Cong-Lem et al. (2025) examined another part of the issue by studying Vietnamese EFL learners' AI use, self-efficacy in using AI, and self-directed learning with AI. They reported a strong positive correlation between AI-use self-efficacy and self-directed learning with AI (r = .71, p < .001). These studies show that AI-supported language learning, learner beliefs, and learner-managed learning have already received attention in the Vietnamese EFL context, but the constructs examined are not identical to those in the present study.

The distinction between these constructs is important. AI self-efficacy concerns learners' confidence in their ability to use AI, whereas speaking self-efficacy concerns their perceived capability to perform English-speaking tasks (Cong-Lem et al., 2025; Li & Zhang, 2023; Phipps, 2024). Similarly, self-directed learning with AI covers learners' management of AI-supported learning more broadly, while AI-SRL in the present study is narrowed to the regulation of AI-supported English-speaking practice (Chang & Sun, 2024; Cong-Lem et al., 2025). A learner may therefore feel capable of using an AI tool without necessarily feeling equally capable of speaking English, and general independent use of AI does not necessarily describe how that learner regulates speaking practice.

Although the academic landscape has begun exploring educational technologies, direct quantitative research addressing the nexus between AI-supported self-regulated learning in speaking practice and speaking self-efficacy remains significantly limited especially within the context of Vietnamese higher education (Cong-Lem et al., 2025; Duong & Suppasetseree, 2024). Extant literature has individually probed adjacent areas such as AI-mediated speaking interventions (Duong & Suppasetseree, 2024), technology-supported self-regulation (Chang & Sun, 2024), L2 speaking self-efficacy (Derakhshan & Fathi, 2024), and broader self-directed AI usage (Cong-Lem et al., 2025); however, a critical research gap persists regarding how these distinct constructs interact in speaking practice. Consequently, the current study seeks to bridge this gap by investigating the extent to which EFL learners' strategic regulation of their AI-supported oral practice predicts their perceived capability to execute English-speaking tasks

The study examines the reported levels of AI-SRL and speaking self-efficacy and the statistical relationship between the two constructs among EFL students at a private university in Vietnam, aligning with emerging research on technology-mediated language learning in the local context (Cong-Lem et al., 2025; Duong & Suppasetseree, 2024). AI utilization is described separately and is not included as part of the AI-SRL composite measure, maintaining the critical theoretical distinction between mere technology usage and active self-regulatory control (Chang & Sun, 2024; Lai & Gu, 2011; Tseng et al., 2006). The four sources of self-efficacy are included only as theoretical background and are not analysed as separate variables (Bandura, 2006). Similarly, the individual phases of self-regulation are not tested as separate predictors within the quantitative model (Zimmerman, 2000, 2002). Finally, because the study employs a cross-sectional correlational design, it does not test whether one construct causes changes in the other, in accordance with established methodological boundaries for non-experimental research (Creswell & Creswell, 2023; Slater & Hasson, 2025). By doing so, these boundaries keep the theoretical framework aligned with the descriptive and correlational analyses used in the present study."

3. Methodology

3.1. Research Design

This study adopted a cross-sectional quantitative correlational design to examine the relationship between AI-supported self-regulated learning (AI-SRL) and speaking self-efficacy among Vietnamese EFL university students. A quantitative approach was appropriate because both constructs were measured through structured questionnaire items and analysed statistically across the participant sample (Creswell & Creswell, 2023). Data were collected at a single point in time, and the correlational design was used to determine the direction, magnitude, and statistical significance of the relationship between the two constructs (Slater & Hasson, 2025).

The study did not manipulate learning conditions or assign participants to experimental and control groups. Its purpose was therefore not to determine whether AI-SRL produced changes in speaking self-efficacy, but to examine whether students who reported stronger regulation of their AI-supported learning also tended to report stronger beliefs in their English-speaking capability. Descriptive statistics were used to address the first research question, while Pearson's product–moment correlation was used to examine the relationship addressed in the second research question.

3.2. Research Setting and Participants

The study was conducted at a private university in Ho Chi Minh City, Vietnam, where English is taught as a foreign language. The institution offers English courses for students who are majors and students who are not majors in English, and the study included students with different academic backgrounds and years of study. The participants were chosen using a convenience sampling method due to its practicality in accessing students who have experience in learning English and practicing speaking with the assistance of AI.

Data were collected over a four-week period from March to April 2026. The questionnaire was administered online through Google Forms and distributed through the university's student communication channels and class groups. A total of 200 responses were initially obtained.

Before statistical analysis, the dataset was screened for incomplete responses and indicators of inattentive responding in accordance with recommendations for detecting low-quality survey data (Ward & Meade, 2023). Particular attention was given to invariant response patterns, or *straight-lining*, in which the same Likert-scale option was selected repeatedly across items that represented conceptually different statements. Such response patterns can indicate limited engagement with questionnaire content. Fourteen responses were excluded because of substantial missing data or clear evidence of invariant responding, leaving 186 valid responses for the final analysis.

The participants also filled out a questionnaire regarding gender, undergraduate degree, academic specialty, prior exposure to AI for language learning and the kind of AI tools they had used. These variables were collected to describe the sample and provide contextual information about participants' experience with AI-supported learning. Not used as independent variables in main correlational analysis.

Because participants were recruited from a single private university through convenience sampling, the sample was not treated as representative of all Vietnamese university EFL learners. The findings were therefore interpreted within the institutional and educational context in which the data were collected.

3.3. Research Instrument

Data were collected using a structured self-report questionnaire consisting of three sections: demographic and AI-use background information, AI-supported self-regulated learning, and speaking self-efficacy.

The first section collected information on participants' gender, year of study, academic major, prior use of AI for English learning, and the types of AI tools they had used. These questions served a descriptive purpose only. They were not combined into an AI-use scale or treated as an additional construct in the inferential analysis.

The second section comprised 13 items measuring AI-supported self-regulated learning. The items were developed or adapted with reference to Zimmerman's (2000, 2002) cyclical model of self-regulated learning and were contextualised for AI-supported English-speaking practice. The initial theoretical structure reflected three related regulatory phases: forethought, performance, and self-reflection.

Forethought items addressed goal setting and strategic planning before speaking practice. Performance items focused on self-monitoring, strategy use, attention, and adjustment during AI-supported practice. Self-reflection items concerned the evaluation of performance, interpretation of feedback, and use of previous practice to inform subsequent learning decisions (Zimmerman, 2000, 2002). This structure allowed AI-SRL to be measured in terms of learners' regulatory behaviour rather than simply their frequency of AI use.

Participants rated each AI-SRL item on a five-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Higher scores indicated stronger reported engagement in AI-supported self-regulated learning.

The third section contained 12 items measuring speaking self-efficacy. The items were developed in accordance with Bandura's (2006) recommendations for constructing domain-specific self-efficacy measures and were contextualised for English-speaking tasks. They assessed learners' perceived capability to express ideas, respond to questions, maintain conversations, communicate despite linguistic difficulties, and perform oral tasks under different conditions.

The speaking self-efficacy items used the same five-point Likert-type response format. Higher scores represented stronger perceived capability to perform English-speaking tasks. Consistent with Bandura's (2006) task-specific conceptualisation of self-efficacy, the scale was designed to measure perceived speaking capability rather than related constructs such as speaking anxiety, willingness to communicate, persistence, or actual speaking proficiency.

Before the main data collection, the questionnaire was piloted with 20 undergraduate students who were not included in the final sample. The pilot was used to examine item clarity, response-format appropriateness, and the approximate time required for completion. Minor revisions were made to improve wording before the questionnaire was distributed to the main sample.

The full item sets were retained at the initial measurement stage. Their psychometric properties were subsequently examined before the final construct scores were calculated.

3.4. Data Collection and Analysis Procedure

Google Forms were used to conduct the questionnaires online. This was self-administered and answers were typed in from the participants. The data was then exported to IBM SPSS Statistics 26 for screening and statistical analysis. The first set of data had 200

answers. Following the above screening process, 14 responses were eliminated, with 186 cases analyzed.

The initial dataset contained 200 responses. After the screening procedure described above, 14 responses were removed, leaving 186 cases for analysis. The remaining data were examined for incomplete responses, coding errors, values outside the permitted response range, and potential response-quality problems. Before Pearson's correlation was calculated, the distributions of the two main variables were also inspected for influential outliers, approximate normality, and linearity because substantial violations of these conditions can affect the interpretation of a Pearson correlation coefficient (Field, 2024).

Descriptive statistics were first calculated to address Research Question 1. Means and standard deviations were used to describe participants' reported levels of AI-SRL and speaking self-efficacy. Demographic information and AI-use background variables were used only to characterise the sample and were not entered into the primary inferential analysis.

Exploratory factor analysis (EFA) was conducted for the AI-SRL scale because the items had been contextualised specifically for AI-supported English-speaking practice. Principal Axis Factoring was used as the extraction method because the purpose was to examine the latent structure underlying the items rather than simply reduce the number of observed variables. Direct Oblimin rotation was selected because the proposed dimensions of self-regulation were theoretically expected to be related rather than independent (Hair et al., 2019).

Before factor extraction, sampling adequacy and factorability were assessed using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity. Factor-retention decisions considered several sources of evidence, including eigenvalues, the scree plot, factor-loading patterns, and the theoretical interpretability of the extracted solution (Hair et al., 2019). Statistical criteria were therefore considered alongside the theoretical structure of SRL rather than used in isolation.

The speaking self-efficacy scale was treated as a single task-specific construct in accordance with the conceptualisation adopted in the study and Bandura's (2006) approach to domain-specific self-efficacy measurement. Its measurement quality was therefore assessed primarily through internal consistency rather than through a separate exploratory factor analysis.

Cronbach's alpha was calculated for both the AI-SRL and speaking self-efficacy scales. Coefficients of .70 or above were regarded as generally acceptable for research purposes (Nunnally & Bernstein, 1994). Corrected item–total correlations and changes in alpha following item deletion were also examined. Item-retention decisions were not based solely on whether deleting an item produced a higher alpha value; theoretical relevance, construct coverage, and consistency with the intended meaning of the scale were also considered.

After factor structure and internal consistency had been evaluated, the retained items for each construct were averaged to produce composite scores. Averaging preserved the original five-point response metric, with higher scores representing stronger AI-SRL or speaking self-efficacy. Construct-level scores were used in the main analyses because the research questions concerned the overall levels of the two constructs and the relationship between them.

Pearson's product–moment correlation coefficient was then calculated to address Research Question 2. The analysis examined the direction, magnitude, and statistical

significance of the relationship between AI-SRL and speaking self-efficacy. A two-tailed test was conducted and a p value less than .05 was considered statistically significant. Cautious interpretation of correlation coefficients was done in consideration of Cohen's (1988) conventional cut-off points of approximately .10, .30 and .50 for small, moderate and large associations, respectively. The values were used as general guidelines for interpreting, not as cut-off points. The results of the coefficient were also judged for their statistical significance in the same direction and with the same size. The study was of a cross sectional correlational design and therefore the correlation was seen as a statistical association, not a cause effect. This distinction determined the way the results were reported and discussed.

3.5. Ethical Considerations

Ethical considerations were incorporated throughout the data-collection process. Before completing the questionnaire, participants received information about the purpose of the study, the voluntary nature of participation, and the confidentiality of their responses. They were also informed that their decision to participate or decline participation would have no effect on their academic standing.

Participants could discontinue the questionnaire at any point before submitting their responses. No personally identifying information was included in the statistical analyses or reported in the study. The data were used solely for research purposes, and all findings were presented in aggregate form to protect participant confidentiality.

4. Results

4.1. Reliability Analysis

Before the main statistical analyses, the reliability of the two adapted scales was examined using Cronbach's alpha. The 13-item AI-Supported Self-Regulated Learning (AI-SRL) scale produced an alpha of .808, while the 12-item Speaking Self-Efficacy scale produced an alpha of .832. Both values indicated acceptable internal consistency, so the two scales were retained for the following analyses.

Table 1

Reliability coefficients of the psychological scales used in this study.

Scale	Number of Items	Cronbach's Alpha
AI-Supported Self-Regulated Learning	13	.808
Speaking Self-Efficacy	12	.832

More detailed reliability results are provided in Appendices A and B. All items in the Speaking Self-Efficacy scale had corrected item-total correlations above .30. For the AI-SRL scale, one item, SRL_F3, showed a lower corrected item-total correlation of .193. Removing this item would have increased Cronbach's alpha only slightly, from .808 to .815. The item was kept because it represented a theoretically relevant part of the AI-SRL construct. Its weaker performance at the item level, though, suggests that it may need further review in future validation studies.

Reliability was not calculated for the AI Utilization section. The items in this section were intended to describe different patterns and purposes of AI use rather than measure a single underlying construct. Calculating one Cronbach's alpha value for these items would therefore not have been suitable.

To examine the internal structure of the adapted AI-SRL scale, an exploratory factor analysis (EFA) was also conducted using Principal Axis Factoring. The Kaiser-Meyer-Olkin (KMO) value was .773, while Bartlett's test of sphericity was statistically significant, $\chi^2(78) = 677.297$, $p < .001$. These results indicated that the data were suitable for factor analysis.

As reported in Table 2, the extracted solution explained 40.04% of the total variance. The observed factor pattern, however, did not clearly reproduce the three proposed dimensions of forethought, performance, and self-reflection.

Table 2

Exploratory factor analysis results for the AI-supported self-regulated learning scale

Statistic	Value
KMO	.773
Bartlett's χ^2	677.297
df	78
Sig.	< .001
Total Variance Explained	40.04%

The EFA therefore offered only preliminary support for the internal structure of the adapted AI-SRL scale. At the same time, the scale showed acceptable internal consistency, with Cronbach's alpha of .808. For this study, the AI-SRL items were therefore retained as a composite measure for the descriptive and correlational analyses. The results involving AI-SRL should still be interpreted with some caution, as the proposed dimensional structure was not clearly reproduced in the EFA. Further validation with larger and more diverse samples would help determine whether the three-dimensional structure can be supported more clearly.

4.2. Descriptive Statistics

Descriptive statistics were calculated to examine the reported levels of AI-supported self-regulated learning and speaking self-efficacy among the participants. As presented in Table 3, the mean score for AI-SRL was 3.65 (SD = 0.53), while the mean score for speaking self-efficacy was 3.62 (SD = 0.57).

Both mean scores were above the midpoint of the five-point response scale. This indicates that participants generally reported favourable levels of self-regulation while using AI and favourable beliefs about their capability to perform English-speaking tasks.

Table 3

Descriptive Statistics for the Main Variables

Variable	N	Min	Max	M	SD
AI-Supported Self-Regulated Learning (AI-SRL)	186	2.08	5.00	3.65	0.53
Speaking Self-Efficacy	186	2.00	5.00	3.62	0.57

AI-SRL produced a slightly higher mean score than speaking self-efficacy. However, the difference between the two descriptive means was small and was not subjected to inferential comparison because the study did not aim to test differences between the constructs. The

speaking self-efficacy score should be interpreted as participants' perceived capability to perform the speaking tasks represented in the questionnaire rather than as evidence of their objectively assessed speaking proficiency.

4.3. Correlation Analysis

Pearson's product–moment correlation analysis was conducted to examine the relationship between AI-supported self-regulated learning and speaking self-efficacy.

As presented in Table 4, a statistically significant moderate positive correlation was found between AI-SRL and speaking self-efficacy, $r = .458$, $p < .001$. This result indicates that participants who reported stronger self-regulation while using AI also tended to report stronger beliefs in their capability to perform English-speaking tasks.

Table 4

Pearson correlation between AI-supported self-regulated learning and speaking self-efficacy

Variables	1	2
1. AI-supported self-regulated learning	—	.458**
2. Speaking self-efficacy	.458**	—

Note. $N = 186$. p < .01, two-tailed.

The result answered the second research question by demonstrating a statistically significant positive association between the two constructs. However, the correlation does not establish causal direction. It cannot be concluded that stronger AI-SRL caused higher speaking self-efficacy or that stronger speaking self-efficacy caused greater self-regulation. The finding indicates only that the two variables were positively associated within the present sample.

5. Findings and Discussion

5.1. Summary of Findings

The present study examined the reported levels of AI-supported self-regulated learning (AI-SRL) and speaking self-efficacy among 186 Vietnamese EFL university students and investigated the relationship between the two constructs. The findings showed that participants generally reported engagement in AI-supported regulatory practices, including goal setting, strategic planning, performance monitoring, feedback evaluation, and self-reflection. They also reported relatively positive beliefs about their capability to perform English-speaking tasks. These results describe how participants perceived their own learning behaviours and speaking capability rather than how effectively they actually regulated their learning or performed in spoken English.

The correlation analysis identified a statistically significant positive relationship between AI-SRL and speaking self-efficacy. Students who reported stronger regulation of their AI-supported learning also tended to perceive themselves as more capable of performing English-speaking tasks. The relationship is therefore meaningful at the level of reported regulatory behaviour and perceived capability within this sample. Its interpretation does not extend to actual speaking proficiency or to changes in either construct over time.

5.2. Discussion of the Findings

Participants generally reported engaging in regulatory behaviours when they used AI to support their English learning. Their responses indicated involvement in planning, performance monitoring, feedback evaluation, reflection, and adjustment of subsequent

learning activities. Rather than portraying AI use as a largely passive activity, this pattern suggests that at least some of the students' reported engagement with AI involved decisions about what to practise, how to monitor progress, and what to do with the information produced during practice.

This pattern can be interpreted through Zimmerman's (2000, 2002) cyclical model of self-regulated learning. In that model, regulation develops through interactions among forethought, performance, and self-reflection rather than through isolated strategy use. The relevance of the model to the present finding lies in the way participants reported using AI across different stages of practice. Planning can give direction to an AI-supported activity, monitoring allows learners to compare ongoing performance with an intended goal, and reflection provides a basis for deciding what should be addressed next. AI-supported practice may provide resources for each of these processes, but the regulatory work remains with the learner.

Research on AI-supported language learning provides related evidence for this interpretation. AI tools can generate practice activities, provide feedback, and give learners opportunities to inspect and revise their language (Kohnke et al., 2023; Su et al., 2023). These technological functions are relevant to the present result because they create opportunities for the types of planning, monitoring, and evaluation reported by participants. The availability of such functions, however, does not establish that learners use them strategically. What distinguishes AI-SRL from AI use more generally is the learner's active evaluation of the information provided and the decisions made on the basis of that information.

The reported level of AI-SRL should therefore be interpreted within the limits of the measurement used in this study. Participants described their own regulatory behaviour; the questionnaire did not directly observe how they interacted with AI or determine whether the feedback they received was accurate. Pintrich (2000) and Zimmerman (2002) both place learner judgement and strategic adaptation at the centre of self-regulation. From this perspective, the present result indicates that participants perceived themselves as engaging in regulatory processes, rather than demonstrating that every instance of their AI use was purposeful or effective.

Reported Level of Speaking Self-Efficacy

Participants also reported relatively positive perceptions of their capability to perform English-speaking tasks. This finding concerns speaking self-efficacy, not objectively assessed speaking proficiency. The distinction matters because students were asked to judge what they believed they could do in speaking situations rather than demonstrate their speaking ability through a performance assessment.

Bandura's (1997, 2006) conceptualisation of self-efficacy provides a useful basis for interpreting this result. Self-efficacy reflects perceived capability for particular tasks and develops partly through the way individuals interpret their previous experiences. Two learners with comparable language proficiency may therefore evaluate their speaking capability differently. In the present context, participants' responses may reflect how they interpreted previous speaking experiences, feedback, perceived progress, and the difficulty of the speaking situations they had encountered.

Bandura (1997) identifies mastery experiences, vicarious experiences, social persuasion, and physiological and affective states as sources from which efficacy information can be derived. These sources offer a theoretical explanation for why speaking self-efficacy may differ across learners even when they study in similar educational settings. A learner who

interprets repeated speaking attempts as evidence of progress, for example, may judge a subsequent task differently from a learner who interprets similar difficulties as evidence of limited capability. The present study did not measure these four sources separately, so they provide an explanatory framework rather than an empirical explanation tested by the current data.

Previous language-learning research also suggests that speaking-related efficacy beliefs are connected with how learners approach oral tasks. Asakereh and Nouraee (2018) linked self-efficacy beliefs with speaking fluency, while Raoofi et al. (2012) examined self-efficacy alongside metacognitive processes and English-speaking performance. These studies offer related evidence rather than a direct comparison with the present finding because the constructs and outcomes examined are not identical. Their relevance lies in showing that learners' judgements of speaking capability form a meaningful part of the broader speaking-learning process.

The participants' reported self-efficacy should consequently not be taken as evidence that they possessed an equivalent level of speaking competence. Bandura (1997, 2006) distinguishes efficacy judgements from actual performance, and the present study maintained that distinction in both measurement and interpretation. What the finding captures is how capable these students believed themselves to be, which is theoretically meaningful but not interchangeable with demonstrated speaking ability.

Relationship Between AI-Supported Self-Regulated Learning and Speaking Self-Efficacy

The principal finding was a statistically significant positive relationship between AI-SRL and speaking self-efficacy. Students who reported stronger planning, monitoring, feedback evaluation, and reflection in their AI-supported learning also tended to report stronger beliefs in their capability to perform English-speaking tasks. The two constructs therefore tended to vary in the same direction within this sample.

Zimmerman's (2000, 2002) model of SRL and Bandura's (1997, 2006) account of self-efficacy offer complementary explanations for why such an association is theoretically plausible. Self-regulated learners establish goals, monitor progress, evaluate outcomes, and use information from previous attempts when deciding how to proceed. Bandura's framework, meanwhile, proposes that efficacy judgements are informed partly by how individuals interpret experiences and performance-related information. The conceptual point of connection lies in this process of interpretation: regulatory activities can make information about progress, difficulty, and strategy effectiveness more available to learners, while efficacy beliefs concern what learners make of their capability in light of such experiences.

Consider a learner who establishes a specific speaking goal and compares several attempts while practising with AI. Monitoring may make gradual changes in pronunciation, fluency, lexical choice, or communicative effectiveness easier for that learner to notice. Feedback evaluation and reflection can then help the learner decide whether those changes represent progress and what requires further attention. Such experiences could become relevant to efficacy judgements when learners interpret them as information about what they are capable of accomplishing. This explanation remains theoretical, since the present study did not directly observe the interpretive process through which individual efficacy judgements developed.

Related language-learning research has connected self-efficacy with performance across various language skills and tasks. Reviewing 32 empirical studies, Raoofi et al. (2012) concluded that self-efficacy was a strong predictor of second- and foreign-language

performance. However, relatively few studies have specifically examined speaking self-efficacy (Phipps, 2024). These studies do not directly test the relationship between AI-SRL and speaking self-efficacy, so they should not be treated as direct empirical confirmation of the present result. Instead, they provide related evidence that capability beliefs can be connected with the regulatory and performance dimensions of language learning. The present study extends the question into the more specific setting of AI-supported speaking practice.

The association may also be understood from the other direction. Bandura (1997) associates stronger efficacy beliefs with goal choice, effort, and persistence, while Zimmerman (2000) places self-efficacy among the motivational processes involved in self-regulated learning. Students who already perceive themselves as more capable speakers may therefore be more willing to establish goals, continue practising when difficulties arise, or engage actively with feedback. The present data cannot distinguish this possibility from the alternative explanation in which stronger regulation is associated with the development of stronger efficacy beliefs.

For that reason, the positive correlation should be interpreted as evidence that AI-SRL and speaking self-efficacy co-occurred systematically within the present sample. It does not identify which construct precedes the other or whether a third factor contributes to both. Longitudinal or experimental research would be needed to examine temporal ordering or causal mechanisms. Within the boundaries of the current design, the finding establishes a statistical relationship and provides a basis for investigating how regulation and perceived speaking capability may interact over time.

5.3. Pedagogical Implications

The findings suggest that pedagogical attention to AI-supported speaking practice should extend beyond access to technological tools. Because stronger reported AI-SRL was associated with stronger speaking self-efficacy, students may benefit from learning how to organize AI-supported practice around explicit goals, monitoring, feedback evaluation, and reflection. This recommendation follows from the observed relationship and the theoretical framework; it should not be interpreted as evidence that SRL instruction will necessarily produce stronger efficacy beliefs.

For students, purposeful AI-supported practice can begin with a clearly defined speaking objective. A learner might focus on fluency, pronunciation, lexical choice, or the ability to respond spontaneously rather than simply asking an AI system to “practice English.” During the activity, learners can monitor the aspect of performance connected with that goal and record difficulties that require further attention. Afterward, they can evaluate the feedback received and decide what should change in the next attempt. This structure keeps decision-making with the learner rather than allowing the technological tool to determine the direction of practice.

Critical evaluation is particularly important because AI feedback is not uniformly reliable. Generative AI may produce inaccurate or contextually inappropriate information, while ASR systems can respond inconsistently to some accents and pronunciation patterns (Kohnke et al., 2023; McCrocklin, 2016). These limitations mean that students need criteria for deciding when feedback is useful and when it should be checked against teacher guidance, task requirements, or other reliable language resources. Developing this judgement is part of learning to use AI strategically rather than merely frequently.

Teachers can support the same process by designing AI-assisted speaking activities around a regulatory cycle. Before practice, students can identify a goal and explain why a

particular AI-supported task is suitable for it. During the activity, attention can be directed toward a limited aspect of performance rather than toward every possible language problem at once. Reflection after the task can then focus on what changed, which feedback was useful, and what should be addressed in the next attempt. This sequence corresponds with the forethought, performance, and self-reflection phases described by Zimmerman (2000, 2002), while still allowing teachers to adapt activities to different speaking objectives.

Frequent AI use should not, in itself, be taken as evidence of strategic learning. Two students may spend comparable amounts of time using the same tool while making very different decisions about goals, feedback, and revision. Classroom tasks can make these decisions visible by asking learners to explain why they selected an activity, which feedback they accepted or rejected, and how they modified subsequent practice. Such requirements place emphasis on the learning process rather than on technology use as an outcome in its own right.

At the curriculum and institutional levels, AI-supported speaking activities can be embedded within structured learning cycles rather than introduced as isolated technology-based tasks. Teacher professional development can similarly address both the pedagogical use of AI and the limitations of AI-generated feedback. Issues such as learner dependence, academic integrity, feedback accuracy, and responsible use remain relevant when AI becomes part of independent language practice. The practical implication of the present findings is therefore not that more AI use is necessarily desirable, but that the quality of learner regulation deserves attention alongside technological access.

5.4. Limitations and Recommendations for Future Research

Several limitations define the scope within which these findings should be interpreted. The first concerns the cross-sectional correlational design. Because both constructs were measured at one point in time, the analysis cannot determine temporal ordering between AI-SRL and speaking self-efficacy. Longitudinal research could examine whether changes in one construct precede changes in the other, while experimental or quasi-experimental designs could test the effects of structured AI-SRL interventions. Such designs would address questions that the present correlational evidence cannot resolve.

A second limitation arises from the exclusive use of self-report questionnaire data. Participants may have overestimated or underestimated both their regulatory behaviour and their speaking capability, and measurement of the two constructs through the same response format may introduce common-method bias. Future research could combine questionnaires with learner journals, interviews, observations, AI interaction records, or learning analytics. Speaking assessments would add another useful source of evidence because they would allow perceived capability to be examined alongside demonstrated performance.

This distinction is especially relevant to speaking self-efficacy. The present study measured what learners believed they could accomplish in English-speaking tasks rather than how proficiently they actually performed those tasks. Future studies could examine AI-SRL, speaking self-efficacy, and independently assessed speaking performance within the same design. Such evidence would make it possible to investigate the extent to which learners' efficacy judgements correspond with observable speaking performance.

Measurement of AI-SRL also warrants further investigation. The scale was adapted specifically for AI-supported English-speaking practice, and although its reliability and internal structure were examined, the exploratory factor analysis did not fully reproduce the proposed multidimensional structure. Future studies could refine individual items and test the revised scale with larger and more diverse samples before using confirmatory factor analysis

to evaluate the proposed structure. Further validation would strengthen the distinction between the theoretical organisation of AI-SRL and its empirical measurement.

The sample places an additional boundary on interpretation. The experiences of all participants may differ from those of learners at public institutions, universities in other parts of Vietnam, and programmes that have different language requirements or technological resources. The research might be multi-institutional to explore the possibility that similar patterns emerge in other disciplines, in other regions, at different proficiency levels, and in different institutions. A wider sampling of data would give better grounds for determining the extent of context in which the observed relationship might be seen. Last, overall level of AI-SRL and speaking SE was analyzed. It did not employ individual SRL phases, Bandura's four sources of self-efficacy, the specific AIs or mediating and moderating variables. Future studies might examine the relationship between goal setting, monitoring, and reflection with speaking self-efficacy and determine if any other variables, including AI competence, digital literacy, language proficiency, language motivation or language anxiety, explain variation in that relationship. This work would not only be a determination of the constructs' association but an attempt to explain when and how they are associated.

6. Conclusion

The present study focused on the level of self-efficacy in speaking and the presence of AI-supported self-regulated learning (AI-SRL) among 186 Vietnamese university students learning English as a foreign language (EFL), and explored the relationship between these two constructs. The participants' self-efficacy in AI-SRL and speaking was relatively high, and the analysis found that there was a significant positive correlation between AI-SRL and speaking. Those students who reported higher levels of consistency across all five categories of goal setting, strategic planning, performance monitoring, evaluating feedback and self-reflection in relation to using AI were also more likely to feel capable of performing English-speaking tasks.

The observed relationship can be understood in the complementary viewpoints of self-regulated learning and self-efficacy.. Zimmerman’s (2000, 2002) cyclical model positions goal setting, monitoring, and reflection as processes through which learners manage and adapt their learning, whereas Bandura’s (1997) social cognitive theory explains how learners form judgements about their capability from interpreted experiences and performance-related information. Considered together, these perspectives provide a theoretical basis for understanding why regulatory behaviour and speaking-related efficacy beliefs may be connected in AI-supported learning. The present evidence establishes this connection at the level of statistical association, not causal influence.

The study adds context-specific evidence from Vietnamese higher education by examining AI-supported self-regulation and speaking self-efficacy within the same empirical framework. Its contribution lies partly in shifting attention away from AI access or frequency of use toward the way learners manage AI-supported practice. From this perspective, educational value is not located in the technology alone. It also relies on whether the learners are able to define goals, self monitor their performance, assess technological feedback and choose further practice. The implication for speaking instruction using artificial intelligence is that it is necessary to implement this interpretation in practice. Giving students access to generative AI or speech technologies alone is not enough, and may not be designed to help them use the technology purposefully. Goal setting, strategic practice, monitoring, critical feedback evaluation, and reflection can then be incorporated into AI-enabled speaking activities for teachers. This kind of support can help to shift the narrative about AI from the idea of a pre-packaged answer to a tool for learners to use.

The conclusions, however, are limited by the design of the study. This study used a cross sectional correlational design, self-reported data, and participants from only one private university and cannot establish the temporal or causal direction between AI-SL and speaking SE. Future studies could build on this research by employing longitudinal and experimental designs, samples from a wider range of institutional settings, and mixed-methods studies that incorporate speaking proficiency along with learner reflections and/or accounts of real-time, AI-supported speaking practice. This evidence would facilitate an understanding not only of the relationship between the two constructs, but also how that relationship changes within learners, across tasks and educational contexts..

References

Asakereh, A., & Dehghannezhad, M. (2015). Student satisfaction with EFL speaking classes: Relating speaking self-efficacy and skills achievement. *Issues in Educational Research, 25*(4), 345–363. https://www.iier.org.au/iier25/asakereh.html

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age Publishing.

Chang, W. L., & Sun, J. C. Y. (2024). Evaluating AI's impact on self-regulated language learning: A systematic review. *System, 126*, Article 103484. https://doi.org/10.1016/j.system.2024.103484

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cong-Lem, N., Nguyen, T. M. P., Nguyen, P. N. N., & Bui, N. V. C. (2025). Exploring Vietnamese EFL learners' generative AI usage patterns, self-efficacy, and self-directed learning: Preliminary findings. In *Embracing voices: Diversity, agency and growth in language education* (pp. 391–408). UEH Publishing House.

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology, 66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Derakhshan, A., & Fathi, J. (2024). Growth mindset, self-efficacy, and self-regulation: A symphony of success in L2 speaking. *System, 123*, Article 103320. https://doi.org/10.1016/j.system.2024.103320

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company.

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, Article 100230. https://doi.org/10.1016/j.caeai.2024.100230

Duong, T. V. T., & Suppasetseree, S. (2024). The effects of an artificial intelligence voice chatbot on improving Vietnamese undergraduate students' English speaking skills. *International Journal of Learning, Teaching and Educational Research, 23*(3), 293–321. https://doi.org/10.26803/ijlter.23.3.15

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121*, Article 103254. https://doi.org/10.1016/j.system.2024.103254

Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th ed.). SAGE.

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology, 27*(2), 6–27. https://doi.org/10.64152/10125/73501

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Hou, F., & Zhou, C. (2025). How AI shapes self-regulated learning in foreign language education. *Discover Computing, 28*, Article 344. https://doi.org/10.1007/s10791-025-09801-0

Kirmizi, O. (2012). Research trends in M.A. ELT programs in Turkey. *Procedia—Social and Behavioral Sciences, 46*, 4687–4691. https://doi.org/10.1016/j.sbspro.2012.06.319

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for language teaching and learning. *ELT Journal, 77*(3), 337–341. https://doi.org/10.1093/elt/ccad024

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lai, C., & Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning, 24*(4), 317–335. https://doi.org/10.1080/09588221.2011.568417

Lai, W. Y. W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. *Computers and Education: Artificial Intelligence, 7*, Article 100291. https://doi.org/10.1016/j.caeai.2024.100291

Li, C., & Zhang, L. J. (2023). The development of accuracy and fluency in second language (L2) speaking related to self-efficacy through online scaffolding: A latent growth curve modeling analysis. *Journal of Psycholinguistic Research, 52*, 1371–1395. https://doi.org/10.1007/s10936-023-09950-7

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), Article 410. https://doi.org/10.3390/educsci13040410

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System, 57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Oxford, R. L. (2011). *Teaching and researching language learning strategies*. Routledge.

Pan, M., Lai, C., & Guo, K. (2025). Technology-enhanced self-regulated second language learning interventions: A systematic review. *System, 131*, Article 103675. https://doi.org/10.1016/j.system.2025.103675

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Phipps, J. (2024). A mixed-method analysis of self-efficacy for speaking with university EFL learners. *The Language Learning Journal, 52*(6), 677–693. https://doi.org/10.1080/09571736.2023.2227866

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press.

Raoofi, S., Tan, B. H., & Chan, S. H. (2012). Self-efficacy in second/foreign language learning contexts. *English Language Teaching, 5*(11), 60–73. https://doi.org/10.5539/elt.v5n11p60

Slater, P. F., & Hasson, F. (2025). Quantitative research designs, hierarchy of evidence and validity. *Journal of Psychiatric and Mental Health Nursing, 32*(3), 656–660. https://doi.org/10.1111/jpm.13135

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology, 14*, Article 1210187. https://doi.org/10.3389/fpsyg.2023.1210187

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78–102. https://doi.org/10.1093/applin/ami046

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System, 127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System, 121*, Article 103259. https://doi.org/10.1016/j.system.2024.103259

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/S15430421TIP4102_2

Bionote

Hoang The Hai Nguyen is currently pursuing a master's degree in English Language at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). His research interests include technology-enhanced language learning, artificial intelligence in English education, self-regulated learning, and EFL speaking development.

Appendix: Questionaire

Title: Examining the Relationship between AI-Supported Self-Regulated Learning and Speaking Self-Efficacy among Vietnamese EFL University Students

Description: This questionnaire aims to explore how students use AI tools in their English speaking practice and how confident they feel about their speaking ability. Your responses are anonymous and will be used for research purposes only. There are no right or wrong answers.

PART A. BACKGROUND INFORMATION

1. Gender:

[ ] Male

[ ] Female

2. Year of study:

[ ] First year

[ ] Second year

[ ] Third year

[ ] Fourth year

3. Major:

[ ] English major

[ ] Non-English major

4. Have you ever used AI tools for learning English?

[ ] Yes

[ ] No

5. Which AI tools have you used for learning English? *(Select all that apply)*

[ ] ChatGPT/Gemini

[ ] ELSA Speak/DuoLingo

[ ] Google Translate

[ ] Others

PART B. AI USAGE

Scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Very often

1. I use AI tools to generate ideas before speaking. [1] [2] [3] [4] [5]
2. I use AI tools to learn vocabulary or expressions for speaking. [1] [2] [3] [4] [5]
3. I use AI tools to practice pronunciation. [1] [2] [3] [4] [5]
4. I use AI tools to improve my fluency. [1] [2] [3] [4] [5]
5. I use AI tools to receive feedback on my speaking. [1] [2] [3] [4] [5]
6. I use AI tools to simulate conversations. [1] [2] [3] [4] [5]

7. I use AI mainly for translation instead of speaking practice. [1] [2] [3] [4] [5]
8. I use AI mainly to get ready-made answers. [1] [2] [3] [4] [5]
9. I use AI tools to repeat speaking practice many times. [1] [2] [3] [4] [5]
10. I use AI tools because they make speaking practice less stressful. [1] [2] [3] [4] [5]

PART C. AI-SUPPORTED SELF-REGULATED LEARNING (SRL)

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree

C1. Forethought – Goal Setting & Planning

1. I set clear speaking goals before using AI tools. [1] [2] [3] [4] [5]
2. I decide which speaking skills I want to improve. [1] [2] [3] [4] [5]
3. I plan how I will practice speaking with AI tools. [1] [2] [3] [4] [5]
4. I choose AI tools based on my learning goals. [1] [2] [3] [4] [5]

C2. Performance – Monitoring & Control

1. I monitor my speaking performance while using AI. [1] [2] [3] [4] [5]
2. I pay attention to AI feedback. [1] [2] [3] [4] [5]
3. I try to correct my mistakes during practice. [1] [2] [3] [4] [5]
4. I adjust my speaking based on AI feedback. [1] [2] [3] [4] [5]
5. I actively interact with AI instead of copying answers. [1] [2] [3] [4] [5]

C3. Self-Reflection – Evaluation & Reflection

1. I evaluate my speaking after using AI tools. [1] [2] [3] [4] [5]
2. I reflect on my mistakes. [1] [2] [3] [4] [5]
3. I think about how to improve in future practice. [1] [2] [3] [4] [5]
4. I compare my progress over time. [1] [2] [3] [4] [5]

PART D. SPEAKING SELF-EFFICACY

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree

D1. Confidence

1. I am confident in my ability to speak English. [1] [2] [3] [4] [5]
2. I can express my ideas clearly in English. [1] [2] [3] [4] [5]
3. I can communicate effectively in conversations. [1] [2] [3] [4] [5]

D2. Willingness to Communicate

1. I am willing to start a conversation in English. [1] [2] [3] [4] [5]
2. I am willing to speak with strangers. [1] [2] [3] [4] [5]
3. I actively participate in speaking activities. [1] [2] [3] [4] [5]

D3. Persistence

1. I can continue speaking even when I make mistakes. [1] [2] [3] [4] [5]

2. I keep trying even when speaking is difficult. [1] [2] [3] [4] [5]

3. I do not give up easily when I lack vocabulary. [1] [2] [3] [4] [5]

D4. Anxiety Control

1. I feel confident even when I am nervous. [1] [2] [3] [4] [5]

2. I am not afraid of making mistakes when speaking. [1] [2] [3] [4] [5]

3. I feel comfortable speaking in different situations. [1] [2] [3] [4] [5]

#44 Application of AI in Teaching and Learning Business English at HUFLIT: strengths, weaknesses, and solutions

Nguyen Thi Bach Yen [1]

Abstract

The research explores the effectiveness of applying Artificial Intelligence (AI) to the teaching and learning of Business English at HUFLIT, specifically in the Customer Service course. The writer uses action research through qualitative and quantitative data collection methods, such as, observation, interview, and survey. The participants include 420 juniors majoring in Business English of six Customer Service classes over a 12-week semester. The study findings show that AI enhances motivation, language use, and communication skills for students. Moreover, the findings show the dependence on AI of some students during the semester. In a nutshell, the research concludes that AI contributes to the teaching and learning of Business English courses positively and effectively with the clear instructions from the lecturers and the appropriate use of the students.

Keywords: Artificial Intelligence (AI), application, Business English, strengths, weaknesses, solutions

1. Introduction

The integration of Artificial Intelligence (AI) into education in general and in teaching and learning English specifically has emerged as a global trend these days. The integration of Computer-Assisted Language Learning (CALL) and Artificial Intelligence (AI) into English as a Foreign Language (EFL) instruction represents a pivotal shift in pedagogical paradigms, driven by the imperative to transcend traditional barriers to language acquisition and foster a more interactive, communicative, and learner-centered educational environment. This transition is necessitated by the evolving demands of the digital era, which calls for an educational approach that not only accommodates the diverse learning styles and needs of students but also leverages technological advancements to enhance teaching effectiveness and language learning outcomes (Hui-Chin, 2024). This trend is particularly relevant at the tertiary level, where both students and instructors often possess the foundational digital literacy required to adopt AI-enhanced methodologies effectively.

Vietnam is no exception to this global movement. In line with global trends, Vietnamese universities and colleges have started integrating AI into teaching and learning practices, especially in specialized fields such as Business English. At Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), this shift has been both timely and strategic. Located in Vietnam's economic and technological hub, HUFLIT benefits from strong infrastructural support, well-trained teaching staff, and a dynamic student population—all of which facilitate the effective adoption of AI-driven educational tools. Faculty members at HUFLIT have received institutional training on how to use AI in lesson planning, instruction, and assessment, significantly improving teaching efficiency while reducing the time spent on administrative tasks. Meanwhile, students are encouraged to use AI tools such as Magic School, ChatGPT, and Gemini to support their learning in Business English

[1] *Faculty of Foreign Languages, Ho Chi Minh University of Foreign Languages - Information Technology (HUFLIT); Email: yenntb@huflit.edu.vn*

courses like Customer Service, Advertising and Marketing, Negotiation Skills, and Marketing in Business.

Despite these advantages, several academic pressures have emerged over the past two years that hinder student learning. Most notably, the academic program has been condensed from four years to three and a half, while the academic calendar has been expanded from two semesters to three per year. These structural changes have led to course overload, leaving students with little time to absorb knowledge, reflect critically, or engage meaningfully with the material. Moreover, some lecturers worried about the abuse of AI from students leading to low results in learning. Therefore, not all lecturers in HUFLIT accept using AI in students' learning process. Devasena (2024) used to say that "Educators don't need to worry about artificial intelligence taking over their jobs. While AI is becoming a valuable tool for educational professionals, there are many ways computers just can't replace the human touch in the classroom."

Given these contextual challenges, the application of AI in language education at HUFLIT is not only timely but essential. By encouraging students to use AI tools for creative, task-based assignments—such as writing business role-play scripts, solving customer service problems, or producing promotional materials—educators can foster more meaningful and interactive learning experiences. In doing so, AI serves as both a cognitive assistant and a motivational force, empowering students to take ownership of their learning while adapting to the demands of a compressed academic schedule.

Based on these goals, the following research questions are posed:

1. *To what extent can artificial intelligence support lecturers and students in the Business English major?*
2. *How useful and effective is the application of artificial intelligence in teaching and learning the Customer Service course?*

2. Liteature review

2.1. The Role of AI in Modern Language Education

The advancement of human intelligence is artificial intelligence. One technology with significantly more education than it currently has is artificial intelligence. The most straightforward way to describe artificial intelligence is as theprocess by which computers mimic human intelligence (AI). Every facet of public life is being increasingly impacted by AI. Both the engine and the foundation of social progress are education (Devasena, 2024). In fact, AI appears in all aspects of life, for instance, education, industry, arts, music, films, engineering, medication, etc.

In language education, AI enables adaptive learning by tailoring content to individual proficiency levels, interests, and learning preferences. The application of AI into teaching and learning English is really practical and updated. There are so many AI tools to support teachers and students in teaching and learning process, such as, Grammarly for grammar feedback, ELSA Speak for pronunciation, Magic School for test composing, ChatGPT for scripts writing, etc. These applications encourage student autonomy and provide opportunities to practice English language skills asynchronously. With the rise of AI use in education, it is important for educators to understand how AI can be used to support the teaching and learning of English to those whose first language is not English. Therefore, a review of extant studies is needed to support the academic and practitioner community in understanding how AI is supporting English language educators and pupils (Crompton et al., 2024). Devasena (2024) mentioned in

his study that the methods and algorithms used in AI-powered tools for language learning vary depending on the specific tool. However, some common methods include natural language processing, machine learning, and deep learning.

In conclusion, if teachers and students do not know the ways to use the advantages of AI, their teaching and learning process is not so smooth and effective as the one of those using it. Consequently, artificial intelligence can effectively support classroom instruction and the development of English communication skills. Therefore, the development of artificial intelligence is not only a significant illustration of how science and technology can advance social change but also a useful tool for teaching and English (Devasena, 2024).

2.2. Strengths and weaknesses of AI in English for Specific Purpose (ESP) Instruction

2.2.1. Strengths of AI

In the era of 5.0, AI plays a vital role in English Language Teaching, especially in the domain of English for Specific Purposes (ESP), such as Customer Service course. Interactive English learning is made possible by artificial intelligence. English learning becomes more stereoscopic and visual through the connection and logical analysis of information such as visuals, music, and text in intelligent systems. Devasena indicated that students interact with AI through a manmachine interface, which not only makes the language environment more genuine but also corrects conversation mistakes in real time, allowing students to learn English in a more efficient manner (2024).

AI tools support these needs through personalized and context-relevant learning. ChatGPT, Gemini, for instance, enable learners to simulate dialogues between staff and customers, staff and managers or business partners, then generate the suitable solutions for each problem in customer care field. AI-generated content like articles, videos, and quizzes adapts to students' interests, lightening teachers' workloads, allowing for personalized support. AI technology revolutionizes English teaching, offering tailored, engaging, and effective learning experiences, adaptable to each student's preferences and needs (Devasena, 2024) . In general, AI-powered tools can be a valuable addition to the teaching toolkit, but they should not be seen as a replacement for human instructors. AI tools can be used to personalize the learning experience, provide timely feedback, and reach a large number of students. However, Shahidi (2023) said that AI tools are not able to provide the same level of in-depth instruction and engagement as human instructors.

In summary, AI is transforming the teaching and learning of Business English, especially in specialized courses like Customer Service. Though issues related to teacher readiness, ethical considerations, and learner dependence persist, the pedagogical benefits of AI—greater engagement, tailored feedback, and real-world skill development—are substantial. Fitria mentioned in his article that "To realize AI's full potential, sustained investment in teacher training, infrastructure, and ethical policy-making will be essential. English education need not be replaced by the Fourth Revolution. Instead of that the development of an English class model using Artificial Intelligence should work in collaboration with the English teaching and learning process. Language literacy and digital literacy are a neat combination to improve global competence" (2021). In the courses like Customer Service in which students must play the roles of service staff, sales manager, business partners, and customers AI tools are really useful because they play role as a peer or assessor. AI can accurately transcribe and assess students' spoken language. Rukiati et al. (2023) mentioned that by comparing their pronunciation, intonation, and fluency to native speakers, AI can provide instant feedback and

suggestions for improvement. Moreover, as virtual language tutors, AI can take part in interactive talks with students.

2.2.2. Weaknesses of AI

However, AI still has some backdraws if students depend on it so much. While AI can significantly enhance the process of second language acquisition, it is essential to note that it should complement, not replace, human instructors and real-life language practice. Songsiengchai et al. showed that AI can provide valuable support and resources, but meaningful communication and cultural understanding often require human interaction. The correlation between AI and SLA is evolving, with technology playing an increasingly prominent role in language education (2023). According to Shimbergenova (2026), when evaluated through the lens of communicative language teaching, the didactic limitations of AI tools become more evident. Communicative competence encompasses not only grammatical accuracy but also pragmatic appropriateness, discourse management, and sociolinguistic awareness AI tools often prioritize form focused instruction and struggle to address pragmatic competence, discourse coherence, and sociocultural nuances of language use. Moreover, overreliance on automated feedback may reduce learners' critical reflection skills if not mediated by instructors. Li and Meng (2021) indicated that current AI systems, despite advances in natural language processing, remain limited in their ability to engage learners in authentic interaction characterized by negotiation of meaning, turn-taking, and contextual interpretation. Empirical studies indicate that AI-mediated dialogues often rely on scripted or probabilistic responses, which may fail to reflect the complexity and unpredictability of real-life communication. Overemphasis on AI-driven practice risks narrowing language learning to decontextualized drills, thereby marginalizing communicative and intercultural dimensions of language use. Research in applied linguistics consistently demonstrates that meaningful interaction is essential for developing fluency and pragmatic competence. As a result, AI tools should be conceptualized as complementary resources rather than substitutes for communicative classroom practices (Shimbergenova, 2026). One another limitation is the AI reliance of students. Research highlights the risk of learner dependency on automated feedback. While immediacy of feedback is beneficial, excessive reliance on AI-generated corrections may reduce learners' capacity for self-monitoring and metalinguistic reflection. Hyland showed that from a didactic perspective, this necessitates balanced instructional design in which AI feedback is supplemented by reflective activities, peer interaction, and teacher-led discussion. Such integration ensures that learners remain cognitively engaged rather than passively responsive to automated prompts (2019). Another weakness comes from the educators. According to Holmes (2019), research indicated that many educators lack sufficient training to critically evaluate AI tools, interpret their outputs, or recognize their limitations. Without such competence, teachers may inadvertently accept automated feedback as authoritative, despite its potential inaccuracies. Didactic planning must include professional development initiatives that equip teachers with the skills to use AI responsibly, ethically, and pedagogically. Studies emphasize that AI tools yield optimal outcomes when teachers actively mediate their use by selecting appropriate tasks, interpreting automated feedback, and linking digital activities to broader learning objectives. Teachers are uniquely positioned to contextualize linguistic input, address learner affective needs, and promote higher-order thinking skills that AI systems cannot adequately support. Without such pedagogical intervention, AI applications may encourage surface-level learning focused on error correction rather than conceptual understanding (Shimbergenova, 2026).

In general, AI tools have both potisitive and negative impacts on teaching and learning English. Their effectiveness depends on pedagogical alignment, ethical governance, teacher mediation, and contextual appropriateness.

3. Methodology

This study adopted the action research design, which is well-suited for educational settings where practitioners aim to improve teaching and learning practices through reflective cycles of planning, action, observation, and evaluation. The research focuses on integrating artificial intelligence (AI) tools into Business English teaching and learning, specifically in the Customer Service course at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT).

For the lecturer, the intervention involved the application of AI tools such as Magic School, ChatGPT, and Gemini in lesson planning, student assessment, and class activities. For the students, these tools were used to support creativity, problem-solving, and communication skills in line with the learning objectives of Business English.

Participants in the study consisted of 420 third-year students majoring in Business English, enrolled in six different Customer Service classes taught by the researcher herself. These students were selected as at an advanced level in their major, with sufficient foundational knowledge to engage meaningfully with AI-supported learning activities.

Data Collection Methods included the following qualitative and quantitative approaches:

- Observation: The researcher conducted structured observations during classroom sessions, focusing on the interaction between students and AI tools, as well as the changes in engagement and learning outcomes. Students' reactions, attitudes and progress were scored each session by the researcher, using the classroom observation checklist during 12 weeks. *(appendix A)*
- Interviews: Semi-structured interviews were held with ten randomly selected students in the researcher's classes. The four interview questions explored the usability, effectiveness, and perceived impact of AI on teaching and learning activities. The answers of students were recorded by the researcher with a mobile phone. *(appendix B)*
- Survey: A Google Form survey was distributed to all 420 students at the end of the semester by sending a suvey link via Zalo groups of six classes taught by the researcher. The survey consisted of five questions and collected insights into their experiences using AI tools for Business English learning, highlighting perceived advantages, limitations, and areas for improvement. *(appendix C)*

All collected data were analyzed thematically to identify recurring patterns and emerging insights. This process allowed the researcher to draw practical conclusions about the extent to which AI can enhance both teaching quality and student engagement in Business English courses.

4. Results

4.1. Findings From The Observation

Overall, the observation data collected from 12 weeks show that the integration of AI into Customer Service course had some notable outcomes. During 12 weeks, students demonstrated their increased engagement, communicative confidence, and interactions with their classmates. In role-play activities, students used AI to generate realistic conversations to

apply the useful languages in the textbook. In problem-solving situations, AI tools were used to simulate customer questions and generate suitable answers. Basing on the scripts written by AI tools, students adjusted into the conversation suitable with the scenarios and then practiced in class. In short, the observation data reveal that AI fosters a dynamic and efficient learning environment for those possessing low competence of language.

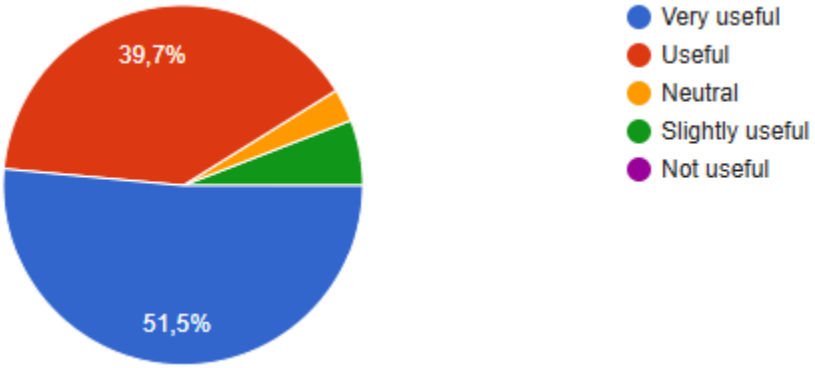
4.2. Findings from the interviews

The qualitative data of the semi-structured interviews including the four open questions indicated that the participants felt happy and satisfied with the integration of AI into teaching and learning Business English. The ten students said that their English speaking skills were improved significantly through some activities, for example, role-plays, problem-solving situations, and group presentations. Furthermore, participants noted that some AI tools lacked contextual sensitivity in business-specific situations. To improve the integration of AI in Business English learning, students suggested more teacher-guided use of AI, integration of AI tools tailored specifically to business contexts, and training on how to critically evaluate AI-generated content. In general, the findings suggest that while AI presents meaningful benefits in Business English education, careful instructional design and digital literacy support are essential to maximize its effectiveness and address potential drawbacks.

4.3. Findings from the survey

Figure 1

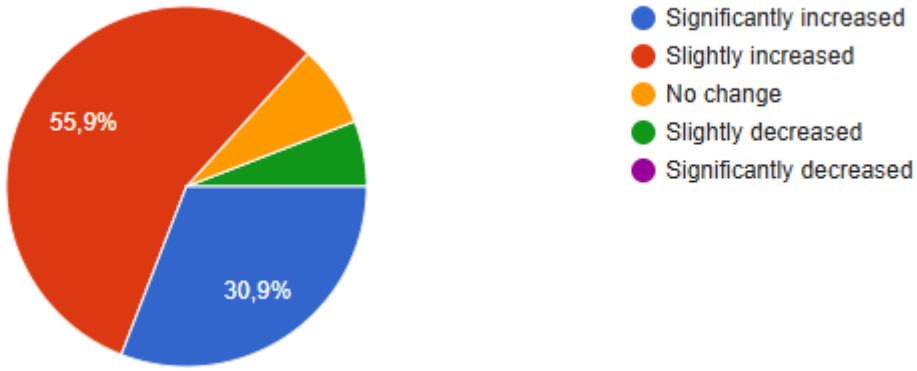
The usefulness of AI tools in Business English learning



The data presented in the figure indicate a predominantly positive perception of AI tools in Business English learning, with 51.5% of respondents rating them as very useful and another 39.7% considering them useful showing that over 90% acknowledge the supportive role of platforms like ChatGPT, Magic School, and Gemini. In contrast, only a small proportion of learners found them slightly useful or neutral, and notably, none deemed them not useful. The absence of negative evaluations further suggests smooth user experience with minimal technological or cognitive barriers. In general, the results highlight the pedagogical potential of AI tools not just as supplementary aids but as integral elements of Business English education, calling for their strategic integration into curricula to enhance engagement, language proficiency, and digital readiness among learners (see Figure 1).

Figure 2

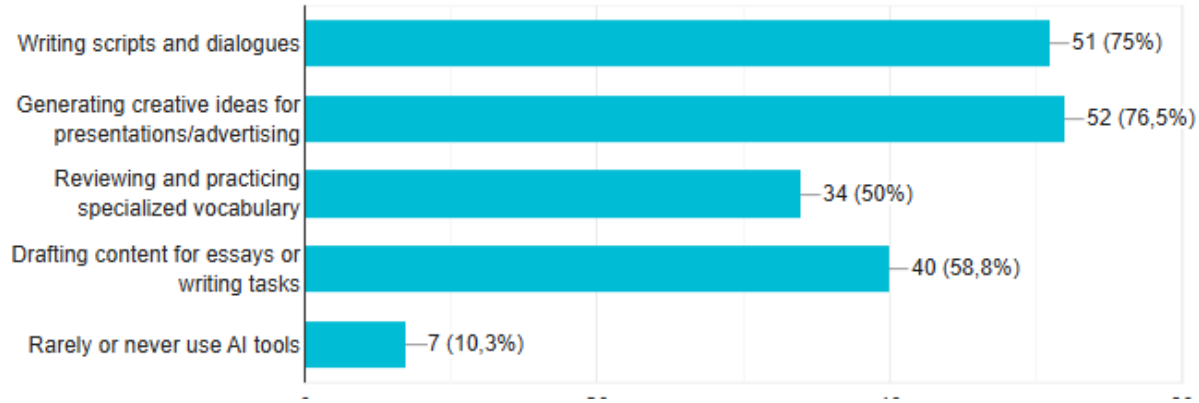
Students' engagement and motivation changes when using AI tools in learning



The results from the survey question "How has your engagement and motivation changed when using AI tools in your learning?" reveal a predominantly positive impact of AI on student motivation. According to the data, 55.9% of respondents indicated a slight increase, while 30.9% reported a significant increase in engagement and motivation when utilizing AI tools. This suggests that a combined 86.8% experienced some form of positive change, highlighting the potential of AI to enhance the learning experience. In contrast, only a small proportion of participants reported either no change (approximately 7%) or a slight decrease (approximately 6%), and notably, none reported a significant decrease. These findings imply that AI tools are generally well-received and can serve as effective motivators in educational contexts, particularly when integrated thoughtfully into learning environments. The minimal negative responses also suggest that concerns about reduced engagement may be less prevalent than anticipated. These results underscore the importance of incorporating AI-enhanced strategies in teaching to support and sustain learner motivation in evolving educational landscapes (see Figure 2).

Figure 3

Types of learning activities students use AI for learning

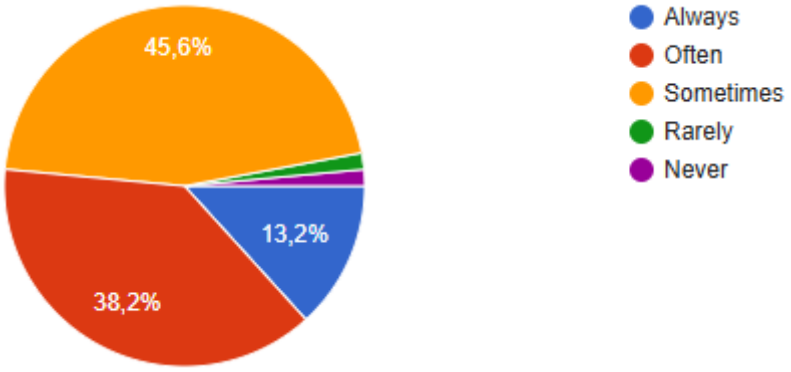


The analysis of the survey results reveals that learners most frequently use AI tools for generating creative ideas for presentations or advertising (76.5%) and writing scripts and dialogues (75%), indicating a strong preference for leveraging AI in creative and communicative tasks. Moderate usage is seen in drafting content for essays or writing tasks (58.8%) and reviewing specialized vocabulary (50%), suggesting AI's role in supporting structured academic and lexical development. Only 10.3% of respondents reported rarely or

never using AI tools, reflecting a widespread integration of AI in language learning practices. These findings imply that AI is increasingly viewed as a valuable support tool in Business English learning, particularly for enhancing creativity, fluency, and domain-specific writing, and highlight the need for educators to incorporate AI meaningfully into instruction while fostering digital literacy and critical engagement with AI-generated content (see Figure 3).

Figure 4

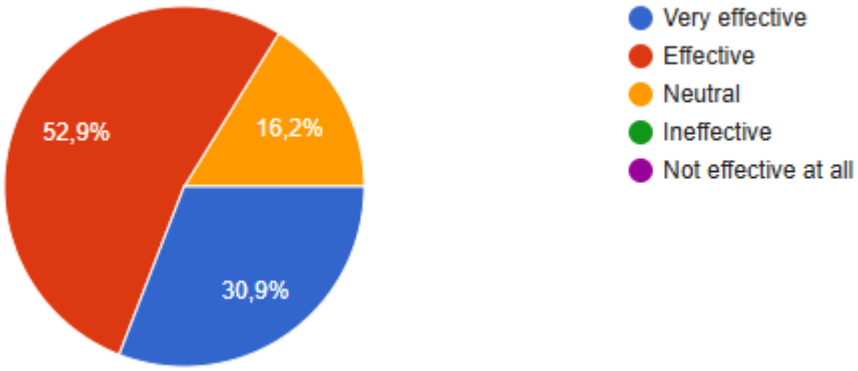
Students rely on AI tools when completing assignments or preparing lessons



The pie chart illustrates the responses to the question, “Do you feel that you rely too much on AI tools when completing assignments or preparing lessons?” The majority of participants selected "Sometimes" (45.6%) and "Often" (38.2%), indicating that a significant portion of respondents perceive a moderate to high level of reliance on AI tools. A smaller proportion chose "Always" (13.2%), while very few reported "Rarely" or "Never," suggesting minimal usage. These findings imply that while AI tools are widely integrated into academic and instructional preparation processes, users are aware of their growing dependence. This suggests a need for further guidance on balanced usage to ensure that AI serves as a support tool rather than a replacement for critical thinking and pedagogical autonomy (see Figure 4).

Figure 5

The overall effectiveness of integrating AI into the Customer Service course



The pie chart presents respondents’ evaluations of the overall effectiveness of integrating AI into the Customer Service course. A majority of participants rated the integration as either "Effective" (52.9%) or "Very effective" (30.9%), suggesting a predominantly positive perception of AI's role in enhancing the course. Meanwhile, 16.2% of respondents expressed a neutral stance, and notably, no participants rated AI as "Ineffective" or "Not effective at all."

These findings imply that the implementation of AI tools in the Customer Service course has been largely well-received, indicating its potential to improve instructional quality, learner engagement, and practical skill development. The absence of negative responses further underscores a strong consensus among learners regarding the benefits of AI integration. Therefore, integrating AI into such vocationally oriented subjects appears to be a promising pedagogical strategy (see Figure 5).

5. Discussion

The integration of Artificial Intelligence (AI) into the Customer Service course at HUFLIT has significantly enriched Business English instruction by fostering student-centered and task-based learning. Platforms such as ChatGPT, ELSA, Magic School, and Gemini facilitated meaningful learner participation in communicative tasks, including role-plays, problem-solving, and script writing. These activities promoted both instructional effectiveness and learner motivation.

In response to research question 1, *To what extent can artificial intelligence support lecturers and students in the Business English major?*, findings indicate that AI can effectively support both lecturers and students by enhancing instructional customization and promoting more autonomous, motivated learning. Survey responses also reflected a degree of concern about potential over-reliance on AI. This underscores the importance of balance: AI should complement, not substitute, traditional instruction. Assessments must be designed to encourage original thought, iterative feedback, and collaborative work. Therefore, teachers must guide their students in adapting AI generated content to appropriate communicative contexts, avoid copy-paste methods.

Regarding research question 2, *How useful and effective is the application of artificial intelligence in teaching and learning the Customer Service course?* the majority of participants found AI both useful and effective for learning Customer Service English, confirming its practical value in ESP education. It demonstrated that with the Fourth Industrial Revolution where technology is very much integrated in teaching and learning, AI is the technology that every educator needs to be exposed to. AI is paramount in the language teaching and learning due to its various benefits (Ali, 2020). The findings will be of interest to educators who intend to integrate AI in their classrooms.

In summary, AI integration in HUFLIT's Business English curriculum has yielded substantial pedagogical gains, including heightened engagement, improved performance, and greater learner independence. However, these outcomes are best achieved when AI is deployed thoughtfully and ethically, under the guidance of skilled educators.

6. Conclusion

This study explores the integration of artificial intelligence in the teaching and learning of Business English at HUFLIT, with a focus on the Customer Service course. Drawing on data from classroom observations, student interviews, and surveys, the research highlights the substantial benefits AI offers in enhancing language instruction, learner engagement, and overall course effectiveness. AI tools such as ChatGPT, ELSA, Magic School, and Gemini were shown to support interactive, student-centered learning experiences by offering real-time feedback, language enhancement, and personalized support. Majority of students showed their motivation and communicative confidence when using AI to do activities such as role-plays, problem-solving scenarios, and group presentations. Moreover, teachers benefited a lot from AI integration through improved classroom management.

However, the research also revealed some concerns about over-reliance and occasional inaccuracies while majority of students appreciated the support AI provided. Therefore, both teachers and students have to combine the traditional pedagogical practices and the advantages of AI so that both can achieve the targets efficiently. Schools must provide students with digital literacy workshops, teachers must give clearer instructions on AI usage to their students before applying it into their teaching and learning at school to have the best results.

In a nutshell, the application of AI in teaching and learning Business English at HUFLIT has indicated to be effective and useful, especially for Customer Service course. When AI continues to develop, teachers and students must remain adaptive and intentional in its usage, making sure that technology serves to empower both teachers and students rather than diminish the autonomy and creation of the learners. Last but not least, future research needs to be done in this field to further explore long-term effects of AI application across different Business English courses at HUFLIT, for instance, Office Communications, English for Business, Business Communications, Negotiation Skills, English for Advertising and Marketing, Selling Skills, etc. and investigate the development of tailored AI tools for professional language education.

References

Ali, Z. (2020). Artificial Intelligence (AI): A Review of its Uses in Language Teaching and Learning. *The 6th International Conference on Software Engineering & Computer Systems.* doi:10.1088/1757-899X/769/1/012043

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology, 55, 2503–2529*. https://doi.org/10.1111/bjet.13460

Devasena R. (2024). Artificial Intelligence in Education: An Alternative to Traditional Learning. *Journal of English Language Teaching, Vol. 66, No. 1.*

Fitria, T., N. (2021). The Use Technology Based On Artificial Intelligence In English Teaching And Learning. *The Journal of English Language Teaching in Foreign Language Context ,Vol 6(2).*

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. *Boston: Center for Curriculum Redesign, pp. 15–22.*

Hui-Chin Yeh. (2024). The synergy of generative AI and inquiry-based learning: transforming the landscape of English teaching and learning. *Interactive Learning Environments.* https://doi.org/10.1080/10494820.2024.2335491

Hyland, K., & Hyland, F. (2019). *Feedback in second language writing*. Language Teaching, 52(3), pp. 335–365.

Li, R., & Meng, Z. (2021). *AI-based pronunciation training*. Computer Assisted Language Learning, 34(5–6), pp. 654–676.

Rukiati, E., Wicaksono, J., A., Taufan, G., T., Suharsono, D., D. (2023). AI on Learning English: Application, Benefit, and Threat. *JLCT, Volume 1, Number 2.*

Shahidi, P., Farsami, M., A. (2023). Teaching Engineering Students English Using Artificial Intelligence (AI) in comparison with Trditional Methods. *8th Iran International Conference on Engineering Education.*

Shimbergenova K. (2026). Didactic Opportunities and Limitations of AI tools in English Language Teaching. *London International Monthly Conference on Multidisciplinary Research and Innovation.*

Songsiengchai, S., Sereerat, B., Watananimitgul, B. (2023). Leveraging Artificial Intelligence (AI): Chat GPT for Effective English Language Learning among Thai Students. *English Language Teaching; Vol. 16, No. 11.*

Bionote

Nguyen Thi Bach Yen is a Doctor of Education (EdD) and a Master of TESOL. She is currently working as a lecturer of English at Faculty of Foreign Languages (FFL), HUFLIT, Vietnam. She likes working in a professional and creative working environment, especially with students who always have new ideas and innovations. With 24 years of teaching, she has been devoting a lot to research.

Appendices

Appendix A: Observation checklist

Observation Criteria	High	Moderate	Low	Notes/Examples
1. Students actively participate in AI-facilitated activities.	☐	☐	☐	
2. Students demonstrate interest/motivation when using AI tools.	☐	☐	☐	
3. Students collaborate with peers during AI-based tasks.	☐	☐	☐	
4. Students ask questions or seek clarification about AI outputs.	☐	☐	☐	
5. Students exhibit improvement or fluency through AI engagement.	☐	☐	☐	
Observation Criteria	**Yes**	**No**	**Partially**	**Notes/Examples**
6. Students use ChatGPT for language practice (e.g., generating responses, role-plays).	☐	☐	☐	
7. Students use Grammarly for writing correction and revision.	☐	☐	☐	
8. Students use ELSA Speak or similar apps for pronunciation practice.	☐	☐	☐	
9. Students engage in negotiation simulations with AI bots.	☐	☐	☐	
10. Students use AI tools independently with minimal instructor support.	☐	☐	☐	
11. Students reflect on AI feedback (e.g., revise based on AI suggestions).	☐	☐	☐	

Aspect	Rating (1–5)	Comments
Integration of AI into the lesson	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5	
Student engagement with AI tools	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5	
Effectiveness of AI in enhancing language learning	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5	

Observation Criteria	High	Moderate	Low	Notes/Examples
Quality of instruction/facilitation	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5			

Appendix B: Interview Questionnaire

Student Interview Questions: AI-Enhanced English Learning

1. How do you usually use AI tools (e.g., ChatGPT, Grammarly, ELSA Speak) when learning Business English?
2. What do you think are the most helpful features or benefits of using AI tools in your Business English learning process?
3. Have you faced any challenges or limitations when using AI tools to learn English? Can you give an example?
4. Do you think using AI tools has improved your English skills (e.g., writing, speaking, vocabulary)? Why or why not?

Appendix C: Survey Questionnaire

SURVEY QUESTIONNAIRE

Title: Application of AI in Teaching and Learning Business English at HUFLIT: strengths, weaknesses, and solutions

Target Group: Third-year Business English majors – Customer Service course

Purpose: To gather students' feedback on the strengths and weaknesses of using AI tools in learning Business English to have suitable solutions

Question 1: How useful do you find AI tools (e.g., ChatGPT, Magic School, Gemini) in supporting your learning in Business English?

☐ Very useful

☐ Useful

☐ Neutral

☐ Slightly useful

☐ Not useful

Question 2: How has your engagement and motivation changed when using AI tools in your learning?

☐ Significantly increased

☐ Slightly increased

☐ No change

☐ Slightly decreased

☐ Significantly decreased

Question 3: What types of learning activities do you most often use AI tools for?

☐ Writing scripts and dialogues

☐ Generating creative ideas for presentations/advertising

☐ Reviewing and practicing specialized vocabulary

☐ Drafting content for essays or writing tasks

☐ Rarely or never use AI tools

Question 4: Do you feel that you rely too much on AI tools when completing assignments or preparing lessons?

☐ Always

☐ Often

☐ Sometimes

☐ Rarely

☐ Never

Question 5: How would you evaluate the overall effectiveness of integrating AI into the Negotiation Skills course?

☐ Very effective

☐ Effective

☐ Neutral

☐ Ineffective

☐ Not effective at all

Thank you for your answers. We will keep them in privacy.

#45 TQM-oriented Service-Learning for EFL-ESL Transition: Social Responsibility at Pham Van Dong University

Võ Duy Quân[1]

Abstract

This study evaluates the efficacy of integrating Total Quality Management (TQM) into Service-Learning (SL) projects to fulfill University Social Responsibility (USR) and facilitate the transition from English as a Foreign Language (EFL) to a simulated English as a Second Language (ESL) environment. In provincial higher education institutions, executing USR is frequently constrained by limited resources, while learners lack authentic language practice environments.

An empirical study is designed at Pham Van Dong University (PDU), involving the entire cohort of third- and fourth-year pre-service English teachers. To optimize resource allocation, the research applies the TQM's PDCA (Plan-Do-Check-Act) cycle and internal quality control criteria to design, operate, and evaluate community-based teaching projects. Data are anticipated to be collected utilizing teaching portfolios, pre- and post-intervention professional agency scales, and quantitative local community social impact metrics.Data analysis is expected to demonstrate a statistically significant increase in pre-service teachers' professional agency post-intervention. Pedagogically, the TQM process is projected to standardize instruction, mitigate deployment risks, and successfully establish a sustainable micro-ESL space.

The findings will provide an empirical foundation for provincial higher education institutions to synchronize USR objectives with language transition policies through rigorous quality management.

Keywords: University social responsibility, EFL-ESL transition, service-learning, total quality management, pre-service teachers.

1. INTRODUCTION

Universities under the jurisdiction of provincial People's Committees have the primary missions of training local human resources, transferring knowledge, and directly addressing local socio-economic issues. Nevertheless, the execution of USR at these institutions faces significant barriers due to a deficiency in fundamental resources. The financial resources of these institutions rely predominantly on provincial budgets and tuition fee revenues. Due to local budgetary constraints, allocated funding is typically prioritized for recurrent operational expenditures. Concurrently, inadequate infrastructure and difficulties in attracting and developing highly qualified faculty members diminish the comprehensive operational efficiency of these institutions.

Consequently, due to resource constraints, provincial universities must often concentrate on pedagogical responsibilities, resulting in insufficient emphasis on the functions of scientific research and community service. As a result, USR has not been systematically integrated into their overall governance strategies. Community activities are predominantly

[1]*Center for Student Support, Corporate Relations and Entrepreneurship, Pham Van Dong University; Email: vdquan@pdu.edu.vn*

campaign-driven and consist of short-term voluntary initiatives conducted by mass organizations. The absence of dedicated personnel and specific funding mechanisms hinders these institutions from implementing academically rigorous and long-term cohesive SL projects. This situation necessitates the restructuring of governance models and the establishment of multi-stakeholder collaboration mechanisms to maintain sustainable USR execution capabilities.

Within the scope of foreign language education, a prominent issue currently is the lack of authentic communicative environments necessary to facilitate the transition of EFL to ESL. At provincial universities, instructional activities predominantly occur within traditional classrooms, focusing on grammar and foundational skills, while lacking contexts for social interaction. This deficiency restricts learners' communicative reflexes and their ability to use the language naturally.

To address this issue, SL is implemented as a central pedagogical approach. Through community projects, such as organizing pre-service teachers to directly teach English to disadvantaged children, learners are transitioned from the classroom environment into practical communicative contexts. Participation in these projects enables pre-service teachers to directly apply their specialized knowledge while simultaneously developing social responsibility, thereby effectively facilitating the transition from EFL to ESL.

Based on this practical rationale, the present study is conducted to evaluate the efficacy of integrating the TQM model into SL projects to fulfill USR and support language practice environments. Rather than pursuing dispersed objectives, the research focuses on two core objectives: (1) Evaluating the operational effectiveness of the PDCA cycle in standardizing pedagogical procedures within SL projects, and (2) Measuring the project's impact on the development of pre-service teachers' professional agency. To achieve these objectives, the study addresses two core research questions:

1. How does the TQM's PDCA cycle assist in standardizing community teaching activities and mitigating risks?
2. How does participation in a TQM-oriented project impact the professional agency of pre-service teachers?

2. LITERATURE REVIEW

2.1. University Social Responsibility and Service-Learning

USR is a foundational governance strategy that directs higher education institutions to create a positive impact on society. This strategy requires institutions to integrate ethical standards and sustainable development goals into two core missions: training and research (Du, 2025; Phan, 2022). To translate the USR strategy into educational practice, SL is utilized as a central pedagogical mechanism. This method establishes a direct linkage between the educational institution's community service mission and the students' practical learning process (Coelho & Menezes, 2021; Lê, 2024).

This linkage operates based on translating governance commitments into practical actions. USR constructs the macro-policy framework, whereas SL serves as the direct implementation tool. The core distinction between SL and volunteer activities is the requirement for students to apply their specialized knowledge to solve local problems, subsequently undertaking steps of academic reflection. This procedure ensures that the educational institution simultaneously achieves its learning outcomes and addresses the practical needs of the community (Nguyễn, 2024; Phạm, 2025; Trần, 2020), while concurrently

developing professional competencies and civic consciousness. The objective of USR is to cultivate a socially responsible workforce. Participation in SL places students in diverse and complex social contexts. The process of resolving practical problems assists learners in cultivating critical thinking, consolidating their professional foundations, and enhancing their awareness of vulnerable groups. Consequently, learners clearly delineate their personal responsibilities towards collective development (Nguyễn & Trần, 2023; Pham & Le, 2025) and establish sustainable partnerships. Through SL, the institution transitions from a unilateral knowledge transfer modality to a bilateral collaborative mechanism with the locality. The educational institution, students, and community partners mutually share resources to address challenges. This interactive mechanism provides empirical evidence affirming the university's capacity for governance and the execution of its social responsibility (Dư, 2025; Ngô, 2023).

2.2. Total Quality Management and the PDCA Cycle in Teacher Education

TQM is a management approach directed towards the continuous improvement of operational processes to meet the standards of the educational system (Bùi & Nguyễn, 2020). The core control instrument of TQM is the Deming cycle, which encompasses four phases: Plan, Do, Check, and Act.

The application of PDCA in teacher education requires educational institutions to transition from management by objectives to process-based control (Nguyễn, 2014). During the Planning phase, the institution determines learning outcomes and program structures based on data from sectoral needs analysis. The Execution (Do) phase involves the organization of instruction, professional training, and the allocation of educational resources. The Checking phase focuses on measuring the actual competencies of learners, as well as collecting feedback data from employers and practicum-hosting institutions. In the Improvement (Act) phase, the institution utilizes the analyzed data to adjust instructional content and training methodologies to prevent procedural discrepancies. The PDCA cycle establishes a self-regulatory mechanism, ensuring that the competencies of pre-service teachers consistently align with practical requirements (Nguyễn, 2023).

2.3. Establishing a simulated ESL environment (micro-ESL space) through Service-Learning

The English language education model in Vietnamese higher education currently operates primarily with an EFL orientation. Students acquire the language within the classroom environment, focusing on grammatical structures rather than utilizing the language as a tool for social communication. To address this issue, SL projects are implemented to establish simulated ESL environments.

According to language socialization theory, language acquisition achieves high efficacy when learners interact within practical contexts. SL provides a natural communicative environment, requiring learners to use authentic language to convey information and execute community-based tasks (Truong et al., 2020). When English-major students directly teach English to children, they are compelled to self-regulate their skills and control their vocabulary to ensure instructional effectiveness. The subsequent process of academic reflection assists students in systematizing their knowledge, rectifying errors, and mitigating psychological anxiety. This constitutes the empirical foundation for students to complete the transition from passive EFL learners to independent ESL users (Bùi, 2023; Tuan & Bao, 2021).

2.4. Professional agency of pre-service teachers in practical environments

Professional agency is the capacity to proactively regulate pedagogical behavior, make independent decisions, and resolve problems within practical contexts (Tran, 2019). For pre-service teachers, this capacity is cultivated only when learners are granted the autonomy to design instructional methods, rather than adhering to pre-existing procedural templates.

The SL environment provides the practical context for pre-service teachers to develop this capacity. Unlike traditional practicums characterized by stringent supervision from mentor teachers, participating in community projects requires pre-service teachers to independently organize educational activities for specific target demographics. Pre-service teachers must directly analyze situations, select pedagogical approaches, and resolve emerging risks (Tuan & Bao, 2021). Through direct practice combined with maintaining reflective journals, pre-service teachers evaluate the consequences of their professional decisions. This process shifts the status of pre-service teachers from passive observers to proactive controllers of pedagogical activities, effectively applying academic theories to authentic educational environments.

3. RESEARCH METHODOLOGY

3.1. Research design and participants

The study adopts a mixed-methods experimental design, conducted at PDU to verify the efficacy of integrating TQM into SL projects.

The research participants comprise third- and fourth-year pre-service English teachers. The quantitative survey sample size is determined to be N = 80 pre-service teachers to ensure reliability and statistical validity.

3.2. Intervention procedure of the PDCA cycle

The experimental intervention is operationalized through the integration of the TQM framework and the institution's internal quality assurance (IQA) criteria, strictly adhering to the PDCA cycle:

Plan: Optimizing the allocation of resources (human, financial, and instructional materials) based on the TQM framework to design projects that closely align with program learning outcomes and local needs.

Do: Organizing pre-service teachers to directly implement practical foreign language teaching activities in community settings.

Check: Utilizing IQA criteria to monitor progress, measure the quality of teaching sessions, and collect feedback from local partners.

Act: Utilizing evaluation data to adjust operational procedures and update pedagogical support methods to standardize subsequent cycles.

3.3. Data collection instruments and methods

Research data are collected concurrently through both quantitative and qualitative channels:

Quantitative instruments: Utilizing the Professional Agency Scale (adapted from the theoretical framework of Tran, 2019), which comprises 18 items measured on a 5-point Likert scale. The survey is administered before and after the intervention (pre-test & post-test) to the 80 pre-service teachers to measure the degree of change. The items are rigorously tested for internal consistency reliability using Cronbach's α coefficient. Concurrently, quantitative social impact metrics (total service hours, number of beneficiaries) are collected to measure the efficacy of USR execution.

Qualitative instruments: Examining Teaching portfolios, including lesson plans and specifically the weekly reflective journals of the 80 pre-service teachers, to evaluate the development in pedagogical thinking and to elucidate the risk management mechanisms underlying the statistical data.

4. RESULTS

4.1. Statistically significant increase in professional agency

To evaluate the impact of the TQM-oriented SL project on pre-service teachers, the study conducted a Paired-Samples T-test on the professional agency scale before and after the intervention. Data were collected from the entire cohort of 80 third- and fourth-year pre-service teachers. Prior to conducting the difference analysis, the data were tested for normal distribution using the Shapiro-Wilk test ($p > .05$), ensuring the fulfillment of conditions for parametric statistics**.**

Table 1. Paired-samples t-test results for pre- and post-intervention Professional Agency (N = 80)

Dimensions of Professional Agency	Phase	M	SD	t	df	p
Overall	Pre-test	3.15	0.62	-15.42	79	**< .001**
	Post-test	4.28	0.45			
1. Pedagogical Awareness	Pre-test	3.20	0.65	-10.85	79	< .001
	Post-test	4.15	0.50			
2. Classroom Management	Pre-test	2.95	0.70	-14.60	79	< .001
	Post-test	4.35	0.48			
3. Social Responsibility	Pre-test	3.30	0.58	-12.35	79	< .001
	Post-test	4.40	0.42			

The data analysis results (Table 1) indicate a statistically significant increase in the pre-service teachers' professional agency post-intervention. Specifically, the overall mean score increased substantially from $M = 3.15$ ($SD = 0.62$) in the pre-test to $M = 4.28$ ($SD = 0.45$) in the post-test ($t = -15.42$, $p < .001$). The most notable improvement was recorded in the Classroom Management dimension (increasing from 2.95 to 4.35), demonstrating that practical exposure robustly assisted pre-service teachers in developing their pedagogical operational and situational problem-solving skills.

4.2. Standardization of instruction and risk mitigation via PDCA

Alongside quantitative data, qualitative data extracted from the pre-service teachers' teaching portfolios clearly demonstrated the role of the PDCA cycle in standardizing instruction and mitigating deployment risks.

4.2.1. Plan and Do

The application of TQM's internal quality assurance IQA criteria right from the lesson planning stage established a robust pedagogical framework. Pre-service teachers were compelled to anticipate local resource constraints to design appropriate lessons. A fourth-year pre-service teacher (SV45) reflected in their journal: *"Initially, I intended to design the lesson based on situational videos and interactive screen-based games. However, upon cross-checking with the resource checklist of the 'Plan' step, I realized that the community cultural*

house was not equipped with a projector system and Internet connection (Wi-Fi). I promptly transitioned to designing large-format flashcards and organizing total physical response (TPR) activities during the 'Do' step. The lesson remained highly engaging; students were not distracted by devices and interacted more directly with the teacher."

4.2.2. Check and Act

The TQM cycle permits the immediate identification of risks rather than waiting until the end of the term. The continuous collection of feedback assists pre-service teachers in adjusting their pedagogical behavior. Another excerpt from a third-year pre-service teacher (SV12) emphasizes the effectiveness of pedagogical risk mitigation: *"After the second teaching session, the 'Check' step with the mentor indicated that I was speaking Vietnamese too much (accounting for 60% of the instructional time). In the following week (Act), I rigorously applied basic Classroom English techniques. Although the local students were initially somewhat unfamiliar, after just two sessions, they responded naturally to English commands."*

4.3. Establishment of micro-ESL space and USR metrics

The integration of TQM into the project not only optimized resource allocation but also successfully established a sustainable simulated ESL environment, directly fulfilling the USR objectives at a provincial higher education institution.

Table 2. *Quantitative local community social impact metrics (USR Metrics)*

USR Metrics	Empirical Data
Resource Allocation	The entire cohort of 80 third- and fourth-year pre-service teachers
Total hours establishing the ESL environment	3,200 hours (80 students x 40 field hours)
Number of direct beneficiaries	520 students/residents in disadvantaged communes
Target Language Use	Averaged 85% of the duration per instructional session
Sustainability Index	100% of community classes completed the committed scheduled duration

The quantitative social impact metrics (Table 2) demonstrate the robust execution of USR. Utilizing a core workforce of 80 pre-service teachers, the project provided 3,200 hours of free instruction, directly benefiting 520 local learners.

Notably, the authentic English utilization rate reaching 85% during the instructional sessions constitutes clear evidence of the successful establishment of a simulated ESL environment (micro-ESL space). Instead of acquiring English through textbooks within a traditional EFL model, local learners were immersed in a simulated ESL environment, where the language is utilized as a dynamic and continuous communicative tool. This affirms that, despite resource constraints, provincial higher education institutions can entirely synchronize their USR objectives with language status transition strategies through a rigorous TQM system.

5. DISCUSSION

This study has demonstrated the efficacy of integrating TQM into SL projects to address specific challenges at provincial higher education institutions. The findings from the empirical

data not only confirm the enhancement of pre-service teachers' competencies but also introduce novel approaches in university governance.

5.1. TQM and Resource Management

At provincial higher education institutions, the execution of USR is frequently hindered by limitations in financial, infrastructural, and human resources. The research results indicate that TQM, particularly the PDCA cycle, functions as a leveraging mechanism to optimize the allocation of these resources.

Instead of necessitating massive investments, the application of IQA criteria from the Planning phase compels the institution and pre-service teachers to accurately identify the existing conditions of the community (e.g., communal cultural spaces, village halls) to restructure them into learning spaces. During the Check and Act phases, the TQM procedure assists in the early detection of pedagogical bottlenecks and resource waste, thereby enabling flexible adjustments without incurring additional costs. In other words, TQM has shifted the governance paradigm from resource deficiency to the optimization of existing assets, transforming SL into a highly feasible and economically viable modality for USR execution.

5.2. USR and Language Transition Policy

One of the strategic contributions of this study is the successful establishment of sustainable simulated ESL environments (micro-ESL spaces). At the macro level, national and local language policies consistently aim to enhance the English proficiency of the general population; however, learners in provincial areas lack authentic language practice environments.

By integrating 80 pre-service teachers deeply into the community, the project created a simulated ESL environment, where English is utilized as a practical communicative tool rather than merely as a pure EFL subject in traditional classrooms. Rigorous quality management serves to integrate two major objectives: executing the institution's community service mission USR while simultaneously establishing an extensive language practice network, thereby directly contributing to the goal of transitioning the status of English locally. This provides a robust empirical foundation for provincial higher education institutions to synchronize their social responsibility objectives with language policies.

5.3. Developing Professional Identity

The statistically significant increase in the professional agency of pre-service teachers post-intervention constitutes crucial evidence for the efficacy of this pedagogical model. Upon participating in the TQM-oriented SL model, the roles of third- and fourth-year students underwent a qualitative shift.

In contrast to traditional practicums where students primarily assume the role of passive apprentices, directly designing, operating, and evaluating community teaching projects based on the PDCA framework requires them to continuously make pedagogical decisions and bear responsibility for output quality. The process of confronting and mitigating deployment risks in field environments has awakened a profound sense of social responsibility. From the position of education beneficiaries, they realize that their professional expertise can address educational inequality issues at the grassroots level. This shift in professional identity, transitioning from a mere language learner to an agent of social change, represents the highest core value that higher education provides to both learners and the community.

6. CONCLUSION AND IMPLICATIONS

6.1. Conclusion

This study has successfully achieved its objective of evaluating the efficacy of integrating TQM into SL projects at PDU. The empirical data affirm that TQM transcends the framework of a business theory to become a rigorous pedagogical management tool. Through strict adherence to the PDCA cycle, the research has successfully standardized instructional procedures, mitigated field deployment risks, and yielded a statistically significant increase in the professional agency of the pre-service teachers.

TQM serves as the pivotal mechanism assisting provincial universities in resolving the challenge of resource scarcity to synchronize two strategic objectives. These are the comprehensive execution of USR concurrently with the successful establishment of a sustainable simulated ESL environment (micro-ESL space), thereby creating momentum for transitioning the status of EFL to a simulated ESL environment for the community.

6.2. Policy Implications

Based on the aforementioned empirical findings, this study provides a data foundation and proposes the following policy implications for provincial higher education institutions:

(1) Restructuring community activities based on Quality Management

Provincial universities need to transition traditional volunteer activities into a SL model oriented by a quality management approach. The integration of IQA criteria into the project design, operation, and evaluation phases is mandatory to optimize resource allocation (human, financial, and instructional materials) and ensure sustainability.

(2) Synchronizing USR with language transition strategies

Local educational administrators should develop mechanisms to integrate social responsibility objectives into foreign language proficiency development schemes. Utilizing the cohort of pre-service teachers as the core workforce to establish authentic language practice environments at the grassroots level not only thoroughly resolves the learners' lack of authentic communicative spaces but also maximizes local budget savings.

(3) Integrating Service-Learning into the Teacher Education curriculum

To develop profound professional agency, it is necessary to formalize the TQM-oriented SL model into a compulsory practicum module or a learning outcome assessment standard for final-year pre-service teachers, rather than merely treating it as short-term extracurricular campaigns.

REFERENCES

Bùi, N. H. V., & Nguyễn, T. N. (2020). Tổng quan lý thuyết và đề xuất mô hình đảm bảo chất lượng giáo dục bên trong trường đại học trong bối cảnh hội nhập và đổi mới giáo dục Việt Nam [Theoretical overview and proposed model for internal quality assurance in universities in the context of integration and educational innovation in Vietnam]. *Tạp chí Khoa học Quản lý Giáo dục, 04*(28), 65–72.

Bùi, T. N. O. (2023). Service-learning project improves essay writing skills of English-major sophomores of Nha Trang University. *TNU Journal of Science and Technology, 229*(04), 364–370.

Coelho, M., & Menezes, I. (2021). University social responsibility, service learning, and students' personal, professional, and civic education. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.617300

Dư, T. C. (2025). Tác động của trách nhiệm xã hội trường đại học (USR) đến tài sản thương hiệu trường đại học theo góc nhìn của sinh viên [The impact of university social responsibility (USR) on university brand equity from students' perspectives]. *HCMCOUJS - Kinh tế và Quản trị Kinh doanh, 20*(4), 76–89.

Lê, N. V. A. (2024). Học tập phục vụ cộng đồng trong giáo dục đại học - Kinh nghiệm thế giới và triển vọng đối với Việt Nam [Service-learning in higher education: Global experiences and prospects for Vietnam]. *Tạp chí Khoa học Trường Đại học Vinh, 53*(Special Issue), 190–199. https://doi.org/10.56824/vujs.2024.htkhgd90

Ngô, T. M. (2023). Trung tâm Service learning trong trường đại học tại Việt Nam trong bối cảnh sáng tạo xã hội [Service learning centers in Vietnamese universities in the context of social innovation]. *Tạp chí Thiết bị giáo dục: Nghiên cứu ứng dụng, 2*(293).

Nguyễn, P. T. (2024). Áp dụng phương pháp lồng ghép Service Learning trong giảng dạy truyền thông [Applying integrated Service Learning in communication teaching]. *Tạp chí Khoa học Trường Đại học Quốc tế Hồng Bàng, 30*, 105–112. https://doi.org/10.59294/HIUJS.30.2024.648

Nguyễn, P. T., & Trần, T. K. (2023). Tiềm năng ứng dụng phương pháp học tập phục vụ cộng đồng (Service Learning) cho đào tạo marketing: Tình huống tại Trường Đại học Kinh tế - Đại học Đà Nẵng [Potential application of Service Learning in marketing training: A case study at University of Economics - The University of Danang]. *Tạp chí Khoa học Kinh tế, 11*(01), 106–122.

Nguyễn, T. H. (2023). Chu trình tổ chức cải tiến chất lượng liên tục "C-EPD" trong đảm bảo chất lượng cơ sở giáo dục [The "C-EPD" continuous quality improvement cycle in educational institution quality assurance]. *Tạp chí Khoa học Giáo dục Việt Nam, 19*(S1), 1–6.

Nguyễn, T. K. (2014). *Quản lý chất lượng đào tạo cử nhân sư phạm theo tiếp cận TQM và ứng dụng công nghệ thông tin trong đại học đa ngành, đa lĩnh vực* [Quality management of bachelor of education training based on TQM approach and IT application in multidisciplinary universities] [Doctoral dissertation, VNU University of Education].

Pham, Q. T., & Le, T. P. L. (2025). Effectiveness of community service learning approach on social responsibility of students of The University of Danang - University of Foreign Language Studies. *Journal of Science and Technology (The University of Danang), 23*(10A). https://doi.org/10.31130/ud-jst.2025.23(10A).459

Phạm, B. T. (2025). Một số vấn đề lí luận về phương pháp học tập phục vụ cộng đồng [Some theoretical issues on service-learning methods]. *Tạp chí Giáo dục, 25*(11), 5–10.

Phan, X. C. (2022). Tác động của trách nhiệm xã hội trường đại học đến thương hiệu và sự gắn bó của nhân viên ở các trường đại học ở TP. Hồ Chí Minh [The impact of university social responsibility on brand and employee engagement in universities in Ho Chi Minh City]. *Tạp chí Khoa học Công nghệ và Thực phẩm*.

Trần, T. B. H. (2020). Xu hướng áp dụng phương pháp học tập phục vụ cộng đồng trong giáo dục đại học, cao đẳng hiện nay và khả năng triển khai tại một số trường khu vực miền Trung - Tây Nguyên [Trends in applying service-learning in current higher education and the feasibility of implementation in some universities in the Central - Central Highlands region]. *Journal of Science – Phu Yen University, 23*, 74–80.

Tran, H. (2019). Teacher agency in times of educational change: The case of transitioned teachers in Vietnam. *Current Issues in Language Planning, 20*(5), 544–559. https://doi.org/10.1080/14664208.2018.1553912

Truong, V., Bui, H. P., & Truong, T. T. L. (2020). Integrating community service learning into university curriculum: Perspectives from EFL teachers and students. *Language Related Research, 11*(5), 201–226. https://doi.org/10.21859/LRR.11.5.201

Tuan, N. H., & Bao, D. (2021). Service learning in a suburban community in Vietnam: Pre-service EFL teachers as agents of change. In D. Bao & T. Pham (Eds.), *Transforming Pedagogies Through Engagement with Learners, Teachers and Communities* (pp. 217–231). Springer. https://doi.org/10.1007/978-981-16-0057-9_14

#46 Examining The Factors Affecting The Employment of ChatGPT in Translation Among English Majors at HUFLIT as Vietnamese Education Transitions From EFL To ESL Orientation

Huong Le[1]

Abstract

The proliferation of ChatGPT has facilitated the learning of translation subjects among language learners, and this is especially true when Vietnam aims to transition English from EFL to ESL. The omnipresence of the chatbot in translation courses proves its effectiveness in assisting learners to achieve their academic targets; nevertheless, the factors behind students' actual employment of the tool seem to have been inadequately explored. This paper aims to investigate the determinants of the employment of ChatGPT in translation among English-majored students. The study involved 138 students of English in their second and third years at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT) utilizing a mixed-methods approach. The findings suggest that intrinsic and extrinsic motivation as well as social influence stand out to be pivotal factors influencing the adoption of ChatGPT in translation assignments among the surveyed students, while the impacts of Perceived Usefulness and Perceived Ease Of Use on the actual use of the tool are negligible. The study also places an emphasis on the equipment of students with essential linguistic knowledge, post-editing competencies, and higher-order thinking skills to effectively utilize ChatGPT for translation subjects while simultaneously encouraging lecturers to make transformations or adaptations in translation pedagogy, drawing students' attention to meaning- and context-oriented translation, fostering interactivity, critical translation and comprehension skills, and ethical awareness of adopting ChatGPT in classroom activities to be aligned with the national orientation of making English become a second language rather than a foreign one.

Key words: factors, employment of ChatGPT, translation, ESL orientation

1. Introduction

The introduction of Artificial Intelligence (AI) tools in language education has made revolutionary changes to both conventional teaching and learning methods (Ghimire et al., 2024), with English being no exception. Teachers have adopted a mixed variety of AI tools to facilitate their teaching process and increase their teaching efficiency (Ghimire et al., 2024), ranging from lesson planning, grading tests, automating assessment (Aljabr & Al-Ahdal, 2024; Amin, 2023; Ferreiro-Santamaria, 2024; Jukiewicz & Wyrwa, 2026; Moqbel & Al-Kadi, 2023; Yavuz et al., 2025) to enhancing the quality of their teaching materials through the creation of stimulating learning activities and the transformation of dry, boring lessons into vivid, intriguing PowerPoint slides with more interactive and appealing content (Li et al., 2025). Meanwhile, learners of language have also taken full advantage of these tools to maximize their learning capacity, undertake self-study, or implement research on the language issues they wish to delve into (Ghimire et al., 2024).

[1] *M.A., Lecturer, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), Ho Chi Minh City, Vietnam; Email: huongltt@huflit.edu.vn*

Translation-related disciplines have also witnessed a dramatic rise in the use of AI tools. Such AI-assisted translation tools as Google Translate, ChatGPT, Gemini, and DeepSeek have played an increasingly important role in enabling learners to handle translation of texts with varying degrees of difficulty. Among these, ChatGPT stands out as a favorable assistant tool for offering immediate, customized feedback tailored to individual learners' demands (Abdallah et al., 2025; Amin, 2023; Ghimire et al., 2024; Pokrivcakova, 2019; Ren, 2025), providing a diversity of linguistic phrases and structures of different levels of complexity, as well as enhancing the precision and speed of translation process (Ren, 2025).

In Vietnam, the employment of ChatGPT in translation learning is increasingly popular, especially when the nation intends to institutionalize English into the education system as a second language across all educational settings by 2035 (Kiet, 2025). Kiet (2025) also adds that the project proposed by the Ministry of Education and Training (MoET) aims to use English as a means of knowledge transmission and day-to-day communication, and this policy shift is to help the country's future labor force successfully integrate into the increasingly competitive global market which is placing growing demands on employees' linguistic abilities. According to the MoET, the new orientation requires English learners to possess better communicative competence, a greater degree of learner autonomy, as well as enhanced lifelong learning capacity and cognitive abilities (Vietnam's smart shift in education, 2025). This also suggests significant implications for translation pedagogy where ChatGPT is mainly adopted out of convenience and personal favorability.

Numerous studies exploring the factors behind the adoption of ChatGPT in educational contexts in general (Al-kfairy, 2024; Alotaibi et al., 2025) and in language settings in particular (Almansour, 2024; Diep & Dang, 2025; Kartini et al., 2025; Qu & Wu, 2024; Ren, 2025) have been conducted since its emergence. However, these studies mainly focused on the indirect effects of such factors as Perceived Usefulness, Perceived Ease Of Use, Motivation, and Social Influence on the adoption of ChatGPT in translation through a mediator called behavioral intention. The direct impact of these factors on the actual employment of ChatGPT in translation in higher education settings has been under-researched. Therefore, this paper seeks to investigate the determinants behind English majors' actual employment of ChatGPT in assisting them in translation courses in a Vietnamese tertiary education institution, specifically in the ESL-transitioned context.

The objective of the study is to examine the factors influencing the employment of ChatGPT in translation among English majors at HUFLIT. The research paper attempts to find out the answers to the following questions:

1. What are the factors influencing English-majored students at HUFLIT to adopt ChatGPT in their translation courses?
2. What skills can be explicitly taught to English-majored students as ChatGPT is formally integrated into translation courses?
3. What kinds of pedagogical instructions or strategies could be incorporated into translation teaching when ChatGPT becomes a formal tool for teaching and learning English?

2. Literature Review

2.1. The use of AI in language education

2.1.1. The benefits of using AI in language education

Plenty of research studies investigating the application of Artificial Intelligence (AI) in language education have credited AI assistant tools for their capabilities of customization, content adaptation to well suit language learners' individual learning needs and the ability to provide detailed analysis of their spoken as well as written outputs and meticulous, instantaneous feedback to help improve their academic performance (Gutiérrez, 2023; Pokrivcakova, 2019; Semerikov et al., 2021; Tafazoli, 2024; Zhu & Wang, 2025). Additionally, AI-powered tools have been found to be able to produce conversations and pieces of writing that bear a great resemblance to those of humans in terms of language use and register (Aljabr & Al-Ahdal, 2024; Gutiérrez, 2023). Most notably, from the perspective of Gutiérrez (2023), AI tools can aid learners in the process of sharpening their linguistic skills owing to its offer of language practice opportunities, along with acting as a source of inspiration (Ali et al., 2023) and engagement enhancement (Abdallah et al., 2025; Ali et al., 2023), which could contribute to interest maintenance among learners. Ferreiro-Santamaria (2024) suggests another important positive effect arising from the adoption of AI in English classrooms, which is the increase in learner autonomy, a prerequisite assisting students in acquiring a language.

2.1.2. The drawbacks of using AI in language education

However, many studies have pointed out some of the limitations in the use of AI tools to facilitate learning process. The first concern involves the production of inaccurate information (Ferreiro-Santamaria, 2024; Gutiérrez, 2023; Vinola et al., 2025), which entails the issue of response reliability (Gutiérrez, 2023; Vinola et al., 2025), and the potential for biases and excessive reliance on content generated by AI (Abdallah et al., 2025; Ferreiro-Santamaria, 2024; Gutiérrez, 2023; Li et al., 2024). The application of AI-powered tools in language education has also been reported to be a major hindrance to the development of teachers' specialized knowledge (Li et al., 2024), the diminution in students' critical thinking (Abdallah et al., 2025; Ferreiro-Santamaria, 2024; Li et al., 2024; Vinola et al., 2025) and social skills due to the shortage of direct interactions, and the increased violation of academic integrity among learners of language (Abdallah et al., 2025; Ferreiro-Santamaria, 2024; Li et al., 2024).

2.2. The adoption of ChatGPT in translation

2.2.1. Translation

According to Ismail & Syahruzah (2017), translation involves the process of conveying meaning of written texts across languages, or more specifically, it is the conversion of meaning from the original language, or commonly known as the source language, into the target language. Sharing the same view; however, Fitria (2018) maintains that translation not only centers around a change in forms, but it also places an emphasis on meaning and style change through the employment of equivalents. A translation of high quality can be produced based on the degree of "accuracy, clarity, and naturalness" (Akil, 2011, p. 2). Among these, the precision of meaning is the most important element contributing to translation quality, which is not just confined to correct idea expression but embraces the transfer of meaning in a contextually and culturally appropriate manner.

2.2.2. The advantages of adopting ChatGPT in translation

On the one hand, the increasingly prevalent use of ChatGPT in translation can be attributed to its competence in generating translations of high degrees of accuracy, pertinence, authenticity, speed with no cost (Roza & Zulhirawati, 2023; Vinola et al., 2025). Time efficiency in proofreading has also been cited as a prominent benefit of using ChatGPT to translate various kinds of texts as students can take advantage of prompts to require ChatGPT to refine the original translations in a diversity of aspects (Vinola et al., 2025). In fact, just by

inputting a specific prompt demanding revision and edition of the translated texts, the tool is able to generate a more precise version of the original one just in less than a second, which helps to save learners a large amount of time on text refinement. Vinola et al. (2025), in their research, have pointed out ChatGPT's flexibility in translation as it can produce texts that are appropriate in terms of styles based on the prompt it receives. An additional strength of using ChatGPT in translation is its brilliant capabilities of structural conformity (Badah et al., 2026), which means the tool is competent in expressing ideas in grammatically correct structures with coherence and without meaning loss.

2.2.3. The disadvantages of adopting ChatGPT in translation

On the other hand, ChatGPT manifests some of its weaknesses, with the first being associated with its ability to translate idioms, cultural subtleties (Manapbayeva et al., 2024; Mohsan & Durr-e-Nayab, 2024; Rousan et al., 2025) and figurative texts as well (Manapbayeva et al., 2024). Because of its limitation in cultural backgrounds, ChatGPT cannot preserve the meaning of the source language when it is translated into the target language. When it comes to translating literary texts, a study implemented by Manapbayeva et al. (2024) reveals that ChatGPT faces a challenge in conveying nuances related to cultures and contexts since its large database tends to provide very general pieces of translation that can overlook cultural sensitivities. Manapbayeva et al. (2024) also raise concerns over the retention of "the stylistic and cultural integrity of the original text" (p. 527) and match these elements with the significance of contextual awareness in translation so as to maintain the purpose and tone of the writer. As for poems, the greatest obstacle to be surmounted by ChatGPT is how to preserve the style and "emotional integrity" (p. 531) conveyed in the original language (Manapbayeva et al., 2024). In their research conclusion, Manapbayeva et al. (2024) suggest the best possible approaches to improve ChatGPT's accuracy in literary translation are to integrate human expertise, feedback, and refinement for better interpretations of cultural and emotional layers existent in the texts. In a similar vein, a recent study undertaken by Albalawi & Abdul-Ghafour (2026) finds that ChatGPT's translations in euphemisms were inconsistently accurate in terms of context, culture, and meaning nuances, thereby resulting in a lack of depth in meaning and cultural inappropriateness. Badah et al. (2026) shares the same view with the previous researchers when pointing out that ChatGPT still struggles with the translation of culturally rich language despite its prominent competence in "structural coherence" (p. 5).

Students who employ ChatGPT as a translation tool also encounter a problem concerning overdependency and reduced translation skills (Vinola et al., 2025). Such skills as researching for suitable and relevant linguistic items and structures, textual analysis, and proofreading, along with the application of linguistic knowledge and cultural consciousness could be greatly undermined if students opt for this AI tool to help them handle translation assignments.

2.3. Theoretical Framework

2.3.1. Technology Acceptance Model (TAM) Theory

A theory of critical importance that serves as a theoretical basis for this study is the Technology Acceptance Model (TAM) theory proposed by Fred David in 1989. Sharp (2007) cited in his research that TAM theory is a model that lays the foundation for the examination of factors influencing users' acceptance of a particular technological tool. Among this model, two fundamental constructs, namely Perceived Usefulness (PU) and Perceived Ease Of Use (PEOU), make an impact on people's attitude towards the use of technology, which in turn influences their behavioral intention of its employment (Alotaibi et al., 2025; Hadalgekar &

Desai, 2025; Lai, 2017). Perceived Usefulness refers to the degree of effectiveness brought about by adopting a new technology on the betterment of an individual's work performance or the accomplishment of a desired result, while Perceived Ease Of Use is defined as the simplicity and user-friendliness of using the technology (Alotaibi et al., 2025; Hadalgekar & Desai, 2025; Harryanto et al., 2018; Lai, 2017; Sharp, 2007). According to Sharp (2007), the findings of the first two studies implemented by Fred Davis in 1989 suggested that Perceived Usefulness exerted a stronger effect on people's adoption intention of a technology compared to Perceived Ease Of Use, and Perceived Ease Of Use preceded Perceived Usefulness, whereas the following study concluded that attitude should be excluded from the model since it did not fully affect the two above-mentioned constructs.

2.3.2. Self-Determination Theory (SDT)

Motivation has been viewed as an element that is of immense importance to the acquisition of a language (Rubić & Matijević, 2019). In other words, success or failure in learning a language greatly depends on how motivated learners are; therefore, gaining an in-depth insight into how motivation impacts on individual learning may assist teachers in adjusting their methods to stimulate students' engagement and interest. Proposed by Deci and Ryan (1985), the Self-Determination Theory is known as a theory examining the relationship between people's motivation and their character traits (Jeno et al., 2017). In this model, students are active learners possessing three fundamental psychological needs, which are "autonomy, competence, and relatedness" (Deci and Ryan, 1985, p. 2, cited in Jeno et al., 2017). These psychological needs are inherent and central to the strengthening of motivation and the direction of people's feelings and behaviors (Rubić & Matijević, 2019). Deci and Ryan (1985, cited in Jeno et al., 2017) categorized motivation into intrinsic and extrinsic motivation, with the former being behavioral actions based on internal sense of delight and pleasure, while the latter indicates behaviors that are exhibited to gain specific outcomes. Annamalai et al. (2023) describe autonomy as the independence to take control of one's own actions or behaviors, defining relatedness as a feeling of connection with other people while referring to competence as the mastery of a task or a related skills set. They elaborate on the correlation between these demands and motivation development and suggest a possible transition from extrinsic to intrinsic motivation once students' psychological needs are satisfied. The Self-Determination Theory also underlines the promotion of a positive learning environment since this can greatly facilitate the nurture of students' psychological needs (Jeno et al., 2017) and highlights the need for instruction to help students regulate extrinsic motivation into internal one, which in turn arouses their genuine curiosity as well as interest and makes them consider learning a rewarding experience (Rubić & Matijević, 2019).

2.3.3. Unified Theory of Acceptance and Use of Technology (UTAUT)

Another theory that is also vital to the understanding of people's decision on adopting technology is the Unified Theory of Acceptance and Use of Technology (UTAUT) formulated by Venkatesh et al. (2003, cited in Venkatesh et al., 2016), which explains the four principal constructs behind individuals' acceptance and employment of a particular technology: Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions (Fatoni & Surani, 2022). Performance Expectancy (PE) is interpreted as individuals' belief of accomplishments in their career performance when putting a system into use, whereas Effort Expectancy (EE) refers to the level of straightforwardness of using the system (Fatoni & Surani, 2022). Social Influence (SI) indicates the impact of significant people on an individual's intention of adopting the system, while Facilitating Conditions (FC) represent the availability of external conditions from an organization that serves as a source of support for

the employment of the system (Fatoni & Surani, 2022). Social pressure can stem from a variety of sources including family members, teachers, peers, and co-workers and has been proved to have a significant influence on people's behavioral intention in academic studies (Ren, 2024).

2.4. Hypotheses

Prior studies conducted on the factors affecting the employment of AI tools in translation show a positive correlation between PEOU, PU, and learners' behavioral intention of adopting these tools (Almogren et al., 2024; Alotaibi et al., 2025; Ren, 2025), which corresponds to the traditional TAM theory. In this research context, behavioral intention construct has been omitted since Vietnamese students have been actively engaged in using ChatGPT for their translation tasks. Based on TAM theory and the context of the research, several hypotheses are formulated as follows:

H1: Perceived Usefulness (PU) positively impacts English majors' actual use of ChatGPT (AU) in translation.

H2: Perceived Ease Of Use (PEOU) positively impacts English majors' actual use of ChatGPT (AU) in translation.

A few studies also found a positive connection between motivation and students' intent of using AI tools in translation (G. C. et al., 2024; Ren, 2025). The following statements are, therefore, hypothesized:

H3: Intrinsic motivation (IM) positively impacts English majors' actual use of ChatGPT (AU) in translation.

H4: Extrinsic motivation (EM) positively impacts English majors' actual use of ChatGPT (AU) in translation.

Social influence, or subjective norms, has been reported to beneficially influence learners' intention of using ChatGPT in language education, according to findings revealed by Almogren et al. (2024), Diep & Dang (2025), and Ren (2025). This entails the formulation of the last hypothesis:

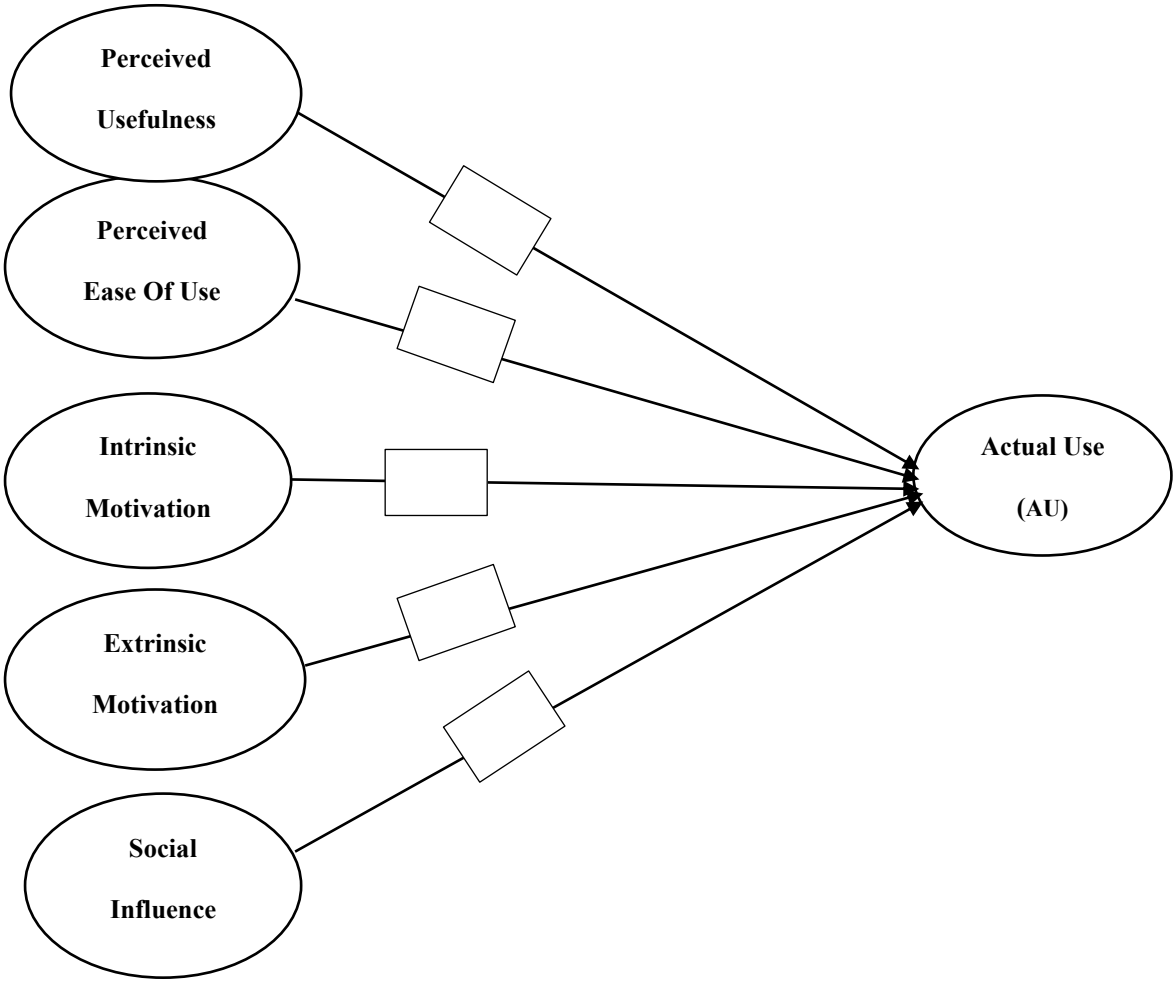
H5: Social influence (SI) positively impacts English majors' actual use of ChatGPT (AU) in translation.

2.5. Conceptual Framework

The conceptual framework is then suggested as follows:

Figure 1

Proposed research model (Source: Researcher's proposal)



3. Methodology

3.1. Participants

The study involved 138 sophomores and juniors, specializing in a variety of majors at Ho Chi Minh University of Foreign Languages – Information Technology (HUFLIT); however, they are all required to take Translation 1 course. Below is the table illustrating the percentages of students in various majors.

Table 1

Participants' demographic information

Majors	Number	Percentage
Translation – Interpretation	15	10.9%
Business English	44	31.8%
English Pedagogy	59	42.8%
English for Office Administration	3	2.2%
English for Logistics	4	2.9%

English - Chinese Bilingualism	13	9.4%

3.2. Instruments

The study was implemented using a questionnaire and structured interviews with lecturers in charge of translation courses and English-majored students. The questions were drafted on Google Forms, and the link to this platform was distributed to groups of students via Zalo, while the interviews were also conducted online right after the completion of the survey. The questionnaire was composed based on such theoretical frameworks as Technology Acceptance Model theory, Self-Determination Theory, and Unified Theory of Acceptance and Use of Technology, and some of the questions were adapted to suit the context of Vietnamese classrooms. The questionnaire consists of nineteen items using a five-point Likert scale (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree) and is divided into diverse themes, namely Background Information, Perceived Usefulness, Perceived Ease Of Use, Intrinsic and Extrinsic Motivation, Social Influence, and ChatGPT Employment in Translation.

The interview sessions were carried out online with the involvement of eight students and four lecturers who have various years of experience in teaching translation courses at HUFLIT. The interviews consist of two questions for students and four questions for lecturers to elicit their perspectives on the employment of ChatGPT in translation teaching and learning.

3.3. Procedures

Subsequent to being drafted on Google Forms, the questionnaire was sent to groups of sophomores and juniors via Zalo through the shortened link obtained on Forms. The purpose of the study was carefully written in the survey to seek informed consent among the participants. Each interview session lasted from approximately fifteen to twenty minutes.

3.4. Data analysis

The data sheet, after being collected through responses on Google Forms, was downloaded and transferred to csv file before being coded to be ready to be imported in SMART-PLS 3.0 software program used for Partial Least Squares Structural Equation Modeling (PLS-SEM), an approach that examines the cause-effect relationship between latent variables, or between independent variables and dependent variables, in a model. These relationships are assessed through the measurement model, the structural model, and the confirmation of hypotheses. Out of 138 responses obtained from Google Forms, 131 responses are valid and used to be imported in the software program.

The interviews with four lecturers and eight students were transcribed and then analyzed using thematic analysis approach following the six steps proposed by Braun and Clarke (2006).

4. Results

4.1. Survey Results

4.1.1. Measurement model validation:

Scale Reliability & Validity Testing:

Table 2

Summary of coefficients in the PLS-SEM model

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
AU	0.818	0.818	0.916	0.846
EM	0.796	0.809	0.880	0.711
IM	0.830	0.840	0.887	0.664
PEOU	0.798	0.803	0.868	0.623
PU	0.853	0.875	0.900	0.692
SI	0.773	0.780	0.867	0.685

Table 3

Outler loadings

	AU	EM	IM	PEOU	PU	SI
AU1	0.922					
AU2	0.917					
EM1		0.876				
EM2		0.773				
EM3		0.876				
IM1			0.875			
IM2			0.817			
IM3			0.835			
IM4			0.725			
PEOU1				0.723		
PEOU2				0.796		
PEOU3				0.803		
PEOU4				0.831		
PU1					0.853	
PU2					0.864	
PU3					0.735	
PU4					0.869	
SI1						0.826
SI2						0.839
SI3						0.818

Cronbach's Alpha (CA) & Outler Loadings: The results of Table 2 show that all construct coefficients are higher than 0.7, specifically ranging from 0.773 to 0.853, which means that all constructs achieve reliability. The findings also illustrate that the outer loadings (Table 3), which are used to assess the quality of the indicators, of all the factors are above 0.7, with the lowest being 0.723. According to Nunnally et al. (1994) and Ringle et al. (2013), these

results guarantee the reliability of all the indicators. Thus, all the variables ensure appropriate scale reliability.

Composite Reliability (CR): Table 2 indicates that the composite reliability of most constructs is greater than 0.8, with that of SI being 0.780, slightly lower than 0.8. Daskalakis & Mantas (2008) suggest when the composite reliability is higher than 0.8, the scale has good internal consistency. However, Henseler & Sarstedt (2012) maintain that CR > 0.7 is acceptable for a model with validation purposes. It can, therefore, be concluded that the internal consistency of the scale has been attained.

Convergent validity (AVE): Hair et al. (2017) prove that when AVE value is greater than 0.5, each construct explains more than half of the variance of its indicator. As can be seen from Table 3, the Average Variance Extracted (AVE) for all the constructs exceeds 0.5, ranging from 0.623 to 0.846, which indicates that each construct achieves good convergent validity.

Discriminant validity (HTMT):

Table 4

Discriminant validity (HTMT) according to the Fornell-Larcker Criterion

	AU	EM	IM	PEOU	PU	SI
AU	**0.920**					
EM	0.663	**0.843**				
IM	0.706	0.689	**0.815**			
PEOU	0.436	0.574	0.586	**0.789**		
PU	0.551	0.555	0.615	0.531	**0.832**	
SI	0.614	0.594	0.661	0.518	0.442	**0.828**

Table 5

Discriminant validity (HTMT) according to the Heterotrait-Monotrait Ratio

	AU	EM	IM	PEOU	PU	SI
AU						
EM	0.817					
IM	**0.853**	0.845				
PEOU	0.533	0.711	0.720			
PU	0.643	0.670	0.716	0.623		
SI	0.759	0.732	0.802	0.645	0.518	

The discriminant validity of the scale was measured using the Fornell-Larcker Criterion and the Heterotrait-Monotrait Ratio (HTMT).

According to the Fornell-Larcker Criterion, the discriminant validity is achieved when the square root of the AVE for any latent variable is higher than any other latent variable's variance. As indicated in Table 4, the discriminant validity of all the constructs is ensured since the square roots of the AVEs (bold and diagonal numbers) are higher than any correlations which are not diagonal.

The data in Table 5 suggest that the HTMT indices for all the pairs of first-order variables for most of the factors are below 0.85, which meets the standard of distinction proposed by Hair et al. (2016). Although the correlation between IM and AU reaches the 0.85 threshold, these are still lower than 0.9, which is an acceptable requirement for discrimination achievement.

Overall, based on the above-mentioned results, it can be said that the scale has ensured good reliability as well as validity.

4.1.2. Structural model assessment:

Multicollinearity Test:

Table 6

Variance Inflation Factor (VIF)

	VIF
AU1	1.916
AU2	1.916
EM1	1.971
EM2	1.461
EM3	1.938
IM1	2.410
IM2	1.832
IM3	1.973
IM4	1.477
PEOU1	1.388
PEOU2	1.513
PEOU3	1.824
PEOU4	2.055
PU1	2.083
PU2	**2.463**
PU3	1.830
PU4	2.260
SI1	1.780
SI2	1.892
SI3	**1.384**

Hair et al. (2017) measured multicollinearity, an issue which refers to the occurrence when there is a high correlation between independent variables, which leads to difficulty in the isolation of their individual impacts on a dependent variable, through Variance Inflation Factor (VIF). If VIF is higher than 0.2 or lower than 5, the independent latent variables do not encounter multicollinearity issues. As shown in Table 6, all the constructs have the VIF of

below 5, with the maximum value of 2.463 and the minimum of 1.384, indicating the absence of multicollinearity in latent variables.

Model Fit Validation:

Table 7

R and R^2 Validation

	R Square	R Square Adjusted
AU	0.594	0.577

The R square value or R^2 value was used to measure the model fit. According to Höck & Ringle (2010), if R^2 value is greater than or equal to 0.67, the explanatory power of the model is considered strong, while R^2 value ranging from 0.33 to 0.67 illustrates a medium power, and if the value is between 0.19 and 0.33, the model has a weak explanatory power. As presented in Table 7, R^2 value of the tested model is 0.577, which means it has a moderate explanatory power.

Table 8

f^2 Validation

	AU	EM	IM	PEOU	PU	SI
AU						
EM	**0.088**					
IM	**0.111**					
PEOU	**0.019**					
PU	**0.030**					
SI	**0.056**					

In this model, f^2 value was used to assess the extent of effect of each construct. Henseler & Sarstedt (2012) conclude that f^2 value is regarded as small, medium, and large if it corresponds to 0.02, 0.15, and 0.35 respectively. Table 8 shows that IM (0.111) has a medium effect, while PEOU (0.019) has a practically negligible impact, with the index being lower than 0.02.

Model Fit

Table 9

Model Fit Validation

	Saturated Model	Estimated Model
SRMR	**0.083**	**0.083**
d_ULS	**1.438**	**1.438**
d_G	**0.638**	**0.638**
Chi-Square	**492.021**	**492.021**
NFI	**0.701**	**0.701**

SRMR (Standardized Root Mean Square Residual) is a goodness-of-fit measure (GoF), which is adopted to measure the quality of the model. If SRMR is smaller than 0.08, the model achieves a good model fit (Hu & Bentler, 1998). The finding reveals that SRMR = 0.083, slightly greater than 0.08 but lower than 0.1, which is an indicator of an acceptable level of model fit.

Table 10

Q^2 Coefficient

	SSO	SSE	Q^2 (=1-SSE/SSO)
AU	262.000	137.631	0.475
EM	393.000	393.000	
IM	524.000	524.000	
PEOU	524.000	524.000	
PU	524.000	524.000	
SI	393.000	393.000	

Table 10 shows that the result of Q^2 coefficient, a value used to measure the level of forecast accuracy, demonstrates a high level of predictive accuracy. There are three levels corresponding to the value of Q^2 coefficient. If the value is between 0 and 0.25, the model has a low predictive relevance. If the value is greater than or equivalent to 0.25 and lower than or equal to 0.5, it is considered medium, and if the value exceeds 0.5, the forecast accuracy level is perceived to be high (Hair et al., 2016; Henseler et al., 2012). As presented in Table 10, the Q^2 coefficient is 0.475, illustrating that the model has a medium level of forecast accuracy.

4.1.3. Research Model Testing and Hypothesis Validation

Figure 2

Estimated model results using SMART PLS-SEM

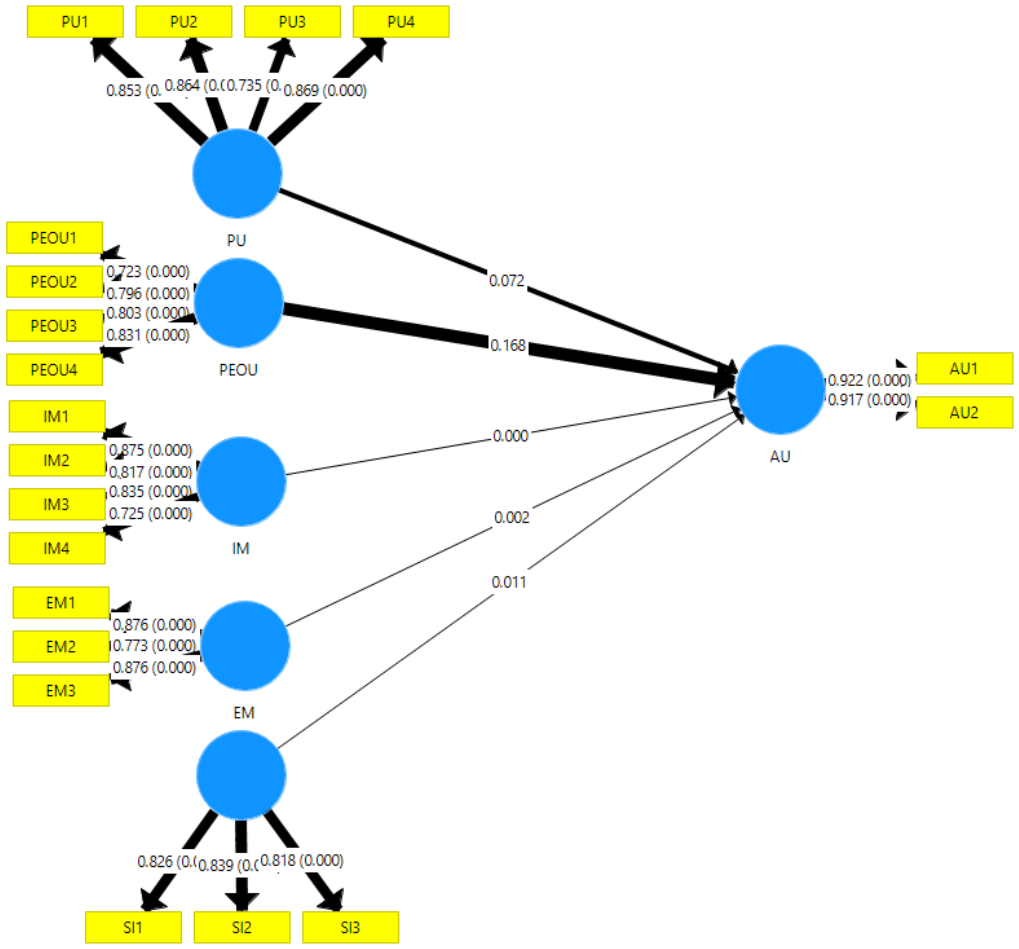


Table 11

Results of hypotheses testing

Hypotheses	Relationships	P Values	Results
H1	PU -> AU	**0.080**	**Rejected**
H2	PEOU -> AU	**0.192**	**Rejected**
H3	IM -> AU	0.000	Accepted
H4	EM -> AU	0.002	Accepted
H5	SI -> AU	0.007	Accepted

Probability value (P value) was employed to evaluate the fit of the structural model. According to Hair et al. (2017), the proposed hypothesis is accepted if P value is smaller than 0.05, and refutation of the hypothesis takes place when P value is greater than or equals to 0.05. Using Bootstrapping 1000, the results indicate that among the five proposed hypotheses, the hypotheses about the effects of IM on AU, EM on AU, and SI on AU are accepted (P values are below 0.05), whereas those of PU on AU and PEOU on AU are rejected (P values exceed 005). This means both intrinsic and extrinsic motivation have a positive impact on the actual employment of ChatGPT in translation, while Perceived Usefulness and Perceived Ease Of

Use exert a marginal or almost no impact on the actual use of this AI tool in textual translation, respectively.

4.2. Interview Results

The interviews with both English majors and lecturers were transcribed and then went through thematic analysis process to extract the most important themes to answer the formulated research questions.

4.2.1. Interviews with students

Three major themes emerged after interviews with eight randomly selected students were held. When asked what skills they should be explicitly taught in translation courses, most of the interviewees agreed on the teaching of linguistic knowledge, higher-order thinking skills, and post-editing competencies.

Theme 1: The consolidation of linguistic knowledge and skills in translation courses

Three out of eight interviewed students underscored the need to consolidate linguistic knowledge and skills for students taking translation courses since it provides a solid foundation for them to evaluate the precision and suitability of grammar and vocabulary use in a ChatGPT-produced translation and confidently correct the errors in the text. They also associated the possession of strong language skills with the capability of assessing the accuracy of meaning conveyed through lexical items and structures in the target language. They seemed to have been more aware of the importance of context and meaning in textual translation rather than the correctness of language use at a superficial level.

"Students need to have good reading comprehension skills and should trust their own abilities instead of relying too much on ChatGPT's translations."

"Students need critical thinking and language skills when using ChatGPT in translation courses. They should be able to check the accuracy of translations and edit them if necessary. For example, a student may use ChatGPT to translate a paragraph, then review the grammar and vocabulary to make sure the meaning is correct."

"*I think students need critical thinking and language knowledge. They should not rely on ChatGPT completely. Instead, they need to check whether the translation is correct and suitable for the context. For example, ChatGPT might translate a word correctly, but the meaning may not fit the situation, so students need to change it."*

Theme 2: The necessity of equipping students with higher-order thinking skills

Higher-order thinking skills are considered crucial skills that students in translation classes should be explicitly taught. This is especially important in the context of Vietnamese education where students have not had many opportunities to hone these skills due to exam-oriented curriculum and teacher-centered approach, to name but a few. Most of the participants suggested training different higher-order thinking skills such as critical thinking, evaluation, reasoning, comparison, researching, and independent thinking skills. Critical thinking, evaluation, and reasoning skills were repetitively mentioned in the students' responses, which indicates that they did not approve of the exclusive reliance on ChatGPT's translation but highlighted the role of learners in analyzing and criticizing the quality of the texts produced by ChatGPT, a process that is laborious and skills demanding. Only by mastering these skills can students evaluate the strengths and weaknesses of the tool, from which they are able to identify areas in need of improvement and make better-informed decisions related to ChatGPT-generated translation versions. Like in this student's response, she cited an example of the

phrase offered by ChatGPT and compared it with another phrase with the same correct meaning but more suitable in the context of Japanese culture. This proves the necessity of developing a critical mind coupled with prior background knowledge application in making lexical choices to convey the meaning of the original text.

"In addition, they should have critical thinking skills and the ability to search for and compare information from different sources in order to produce the best possible translation. For example, when translating the word "hoa đào," ChatGPT may translate it as "peach blossom." However, in some contexts such as Japanese culture, the correct translation should be "cherry blossom" (sakura). Therefore, students need to check the cultural context and compare different sources before deciding which translation is the most appropriate."

In agreement with the previous student, the following interviewee also highly valued the ability to think critically when it comes to translating culturally sensitive elements like idioms as ChatGPT's lack of cultural knowledge can hinder it from expressing the meaning of idioms naturally and accurately.

"I think students need strong critical thinking and post editing skills. We shouldn't just copy paste, instead we must be able to spot errors and refine the AI's output to make it sound more natural. For example, if ChatGPT translates a Vietnamese idiom literally, a student needs the skill to recognize that it sounds weird and change it to a more suitable English expression."

One student answered in the interview that the skill of most crucial importance was independent thinking skills. He asserted that students had better depend on themselves to handle the assignments, which is reflective of his consciousness of learner autonomy enhancement in translation, a requisite for lifelong learning. Self-reliance could not only free language learners from excessive dependence on ChatGPT, but it also helped stimulate their creativity and flexibility in idea construction.

"I think we have a clear example through the name. One is ChatGPT, and one is YOU. For that, the skill I think most required is do not let yourself be bound by a tool."

Theme 3: The significance of post-editing competencies in translation courses in ESL-oriented context

Five out of eight students stated that post-editing competencies should be incorporated into translation courses when they used a variety of lexical items such as "check," "change," "adjust," "edit," "spot errors," and "refine" to refer to the skills. ChatGPT is capable of coping with error detection and correction; however, the students thought they should be those shouldering this responsibility and to achieve this, they had to master the skills of post-edition. The process of post-editing a text requires a combination of competencies relevant to the comprehension and application of diverse linguistic aspects, identification of mistakes, knowledge of lexical and structural equivalents in the target language, and the ability to make prompt decisions in textual refinement, which can be nurtured through proper training.

"I think that students must have the ability to evaluate and reason when receiving information. For example, when receiving translation results, I would check and evaluate their accuracy and grammar. From there, I would adjust them to suit the target language."

Prompt engineering is also suggested as a necessary skill for honing as ChatGPT is formally used in translation courses. The student expressing this viewpoint said that prompts should be specific and precise to improve the tool's competence in generating translation versions that match learners' expectations. This response illustrates his recognition of the tool's limitations and the awareness of context and style attached to the texts.

"For example, instead of just saying "Translate this," a student should be able to command: "Translate this paragraph into academic English, using passive voice to enhance objectivity."

4.2.2. Interviews with lecturers

With regard to the question of how the adoption of ChatGPT has changed the way they teach translation subjects, the lecturers had differing opinions which could be divided into three themes.

Theme 1: Making transformations or adaptations to translation teaching.

Mindset change and role switching

Three out of four lecturers acknowledged that they had to transform their translation pedagogy once ChatGPT is formally incorporated into the curriculum, starting with a change in attitude towards the employment of ChatGPT in translation. They said that ChatGPT should be accepted as an integral part of the teaching and learning process instead of being banned from use; however, students should receive guidance on how to use ChatGPT appropriately. This change in mindset could be beneficial to students' linguistic development due to more frequent exposure to correct English, which may also improve their decision-making competency.

"My point of view is that in the current AI boom, instructors cannot and should not prohibit students from using ChatGPT to support their learning. Instead of prohibiting them, instructors should allow them, especially those studying Translation, to use apps they find necessary and effective for their learning."

Another transformation involves a shift in lecturers' role from those who traditionally act as knowledge providers to facilitators whose responsibility is to support and instruct students to critically evaluate ChatGPT-produced translations. Classroom activities, therefore, center around comparison and analysis of differences in linguistic and cultural elements between translations written by students and those produced from the chatbot.

"Instead of focusing only on traditional translation methods, I now integrate AI into classroom activities. For example, I ask students to compare their own translations with ChatGPT's output, then analyze differences in accuracy, tone, and cultural appropriateness. This has shifted my role from being the sole provider of answers to more of a facilitator who guides students in critical evaluation."

A respondent suggested that lecturers were expected to take on the role of linguistic scaffolders and monitors. To discourage students from excessive reliance on ChatGPT, lecturers should be more supportive in providing useful words, phrases, structures, collocations, and kindred issues while monitoring students' understanding and engagement in translation exercises in the classroom.

"Translation assignments usually have topics that align with the expectations of the course syllabus. Therefore, I will require students to grasp the topic first. Then, I divide the text into short paragraphs. Students will go to the board, only allowed to bring their textbooks, and translate the requested paragraphs. Students are not allowed to bring phones or other devices to the board. While one student is translating on the board, as an instructor, I suggest equivalent vocabulary or synonyms and especially equivalent grammatical structures to the class so that other students can create a similar or equivalent version. While students are translating on the board, the instructor checks their vocabulary and grammatical structures to

check their understanding of the text. For example, why is there ed in this phrase; why is which needed here; why isn't this sentence separated into two parts, etc."

Some of the lecturers also expressed their concerns over adaptations that should be made to translation teaching when English is turned into a second language at tertiary level. One lecturer felt that the integration of ChatGPT has enabled her to prioritize aspects like problem-solving and reflection in her lesson design process by encouraging her students to identify ChatGPT's areas of struggle in translation, particularly with idioms and specialist vocabulary.

"It also allows me to design lessons that emphasize problem-solving and reflection, such as identifying idioms or technical terms that AI struggles with."

Another interviewee admitted that she had to make use of blended teaching approach in her class, developing students' learner autonomy by giving them assignments for pre-class practice and requiring them to do in-class translation practice with the absence of AI tools.

"I have to change many activities which can help them remember the meaning and usage of vocabulary and structures in both languages. I have them translate the passages from the textbook at home before class time. I usually ask them to practice translating in class without AI tool or dictionary."

Designing more translation activities that promote interactivity, focus on meaning and context, as well as fostering students' critical translation and comprehension skills

Translation is usually considered a monotonous activity, so the incorporation of ChatGPT into the curriculum, from a lecturer's viewpoint, is likely to help increase interactivity among students in the classroom. According to another interviewee, games should also be introduced to boost memorization of new lexical items and structures as well as increasing interaction among learners.

Most of the interviewees agreed on the enhancement of students' critical thinking, decision-making, and evaluation skills through the organization of comparison, error analysis, reflection journals, and collaborative or group-based projects. Two lecturers stated that comparison exercises would likely benefit students in terms of linguistic accuracy, style development, and cultural awareness. Error analysis, reflection journals, and groupwork projects would sharpen students' evaluation and post-editing skills, reflecting on their improvements, and helping them value progress more than product.

"*One useful activity is comparison tasks, where students translate a text themselves and then compare it with ChatGPT's version, discussing differences in accuracy, tone, and cultural meaning. Another strategy is error analysis, where students identify mistakes in AI translations, such as idioms or technical terms, and propose better alternatives. Teachers can also design collaborative projects, asking groups to refine ChatGPT's draft into a polished translation, which encourages teamwork and critical thinking. Additionally, reflection journals can be used, where students record how they used AI, what they learned, and how they improved the translation. These activities help students see AI as a supportive tool, while still developing their own judgment and translation skills."*

A lecturer thought that translation lessons should encompass more activities devoted to sentence meaning and register rather than grammatical structures, which could assist learners to distinguish human translations from those produced by ChatGPT. It is the emphasis on meaning that can promote learners' recognition of their own importance in making their translations more easily distinguishable from machine-generated texts.

"... Lessons should not only focus on structures but the core meaning of sentences, and they should have a more focus on style of translation, which I think can differentiate human and machine's use of language and structures."

Despite this, one lecturer emphasized the importance of linguistic knowledge building, pointing out activities focusing on comprehension and translation skills such as figuring out, applying synonyms, opposites, collocations in contexts, and answering guided questions had better be consolidated during class time.

"[...] find synonyms, find antonyms, find equivalent structures, find collocations of keywords in the lesson, put the collocations in contexts or sentences. At the beginning of the lesson, the teacher asks students about vocabulary learned in the previous session."

She specified that these activities could provide a solid foundation for students to gain a deeper understanding of word and structure usage, thereby enhancing their comprehension skills.

"When the student can answer the instructor's questions, the AI-based reference is considered effective. Conversely, if the student only knows how to write the translation on the board without understanding the instructor's why questions, their AI-based look-up is considered rote learning and ineffective reference."

Theme 2: Skills that students of translation courses should be equipped with

The four lecturers who attended the interviews suggested a diversity of skills on which students should receive explicit instruction in translation classes. They perceived the foundational skills students should master to be lower-thinking and translation skills, insisting that students had better remember the vocabulary and grammar patterns acquired and should be able to apply them in sentences. They believed learners should also be provided training on how to effectively utilize such strategies as thematic or topical understanding, paraphrasing, simplification of complicated ideas, appropriate employment of register in translation during the courses. Other skills they mentioned are higher-order thinking skills, enhanced awareness of context and cultural nuances, and post-editing competence. They maintained that critical evaluation ability for textual precision and meaning allows students to assess the translated texts objectively before making any relevant adjustments, whereas detecting ChatGPT's weaknesses in literary translation and cultural subtleties helps improve learners' judgement, and post-editing competency can result in the best possible translation versions. Crafting prompts, from the perspectives of two lecturers, is also an essential skill that will likely aid students in requiring ChatGPT to generate much better results tailored to their requirements.

"The following can be explicitly taught as related skills: how to grasp the theme of the text, how to translate concisely, how to create equivalent structures, how to rephrase or simplify complex ideas, how to put the collocations in contexts or sentences, how to improve formality using appropriate styles or nuances in different contexts, and how to have a better version."

"Students can be explicitly taught several key skills when ChatGPT is formally integrated into translation courses. First, they need critical evaluation skills, learning how to check AI translations against the source text for accuracy and meaning. Second, editing and revision skills are important, so students can refine AI outputs to match academic or professional standards. Third, they should practice cultural and contextual awareness, recognizing when AI struggles with idioms, metaphors, or culturally specific expressions. Finally, prompting skills can be taught, showing students how to give clear instructions to AI

for better results. For example, a student might ask ChatGPT to "translate this text into formal academic English" rather than just "translate," which produces more precise outcomes. These skills ensure students use AI responsibly while still developing their own translation competence."

Theme 3: The types of pedagogical instructions or strategies that could be incorporated into translation teaching when ChatGPT becomes a formal tool for teaching and learning English

Pedagogical instructions on ChatGPT adoption

All the lecturers held the belief that students should be provided with pedagogical guidance on AI adoption. On the one hand, students should be officially informed of the acceptance of ChatGPT as a supportive tool in the classroom, which means the tool can be utilized to improve their linguistic competence. On the other hand, they had to conform with some regulations in tests and assessment where AI tools were restricted, and they were expected to maintain self-reliance in exams. Furthermore, students should be instructed to write prompts more effectively to meet their demands and the objectives of the unit. This kind of AI literacy training can help students gain deeper insights into their translation tasks, especially in terms of specialized words, styles, and cultural variations, while cultivating a sense of critical thinking, analytical judgement, along with learner autonomy.

"*When my students intend to use ChatGPT in their translation exercises, I usually give them clear instructions to guide their practice. First, I remind them to use ChatGPT as a support tool, not as a substitute for their own thinking."*

"Announcing that students are allowed to use translation tools when doing assignments such as classwork, homework, Moodle assignments, and emphasizing that in the mid-term and final exams, students must rely on themselves. Clarifying that any assistance from translation applications or tools used in the exam room will be considered cheating. Informing learners that translation tools are merely supporting tools, while learners themselves must learn various translation methods to create as many versions as possible, with the ultimate goal of boosting translation skills."

Instructions on ethical use of ChatGPT

The second type of instruction suggested by the lecturers is raising ethical awareness in adopting ChatGPT among learners by specifying that they were supposed to use ChatGPT in a responsible manner and be conscious of biases and other sensitive issues arising from the employment of this AI-assistant tool since increased awareness of these problems can prevent them from exhibiting a distorted view of other cultures or languages.

"In the middle of the course, I train my students how to use AI tools effectively and discuss with them about ethical issues. Specially, they need to evaluate the AI translation and be responsible for their final translation version."

"... I mean, they'd be able to manually check for accuracy and identify biased results, and beware of several ethical issues such as privacy, security and sensitivity of generated data."

Instructions on development of learner autonomy

Despite permitting the employment of ChatGPT for linguistic practice, the lecturers expected their students to be actively and critically engaged in translation tasks through comparison and contrast of their own translations with those of ChatGPT, accuracy check, and

revision of ChatGPT-generated versions. In addition, they ought to be involved in decision-making process and take advantage of translation methods and skills to produce multiple translations. This implies the prioritization of enhancing translation competence and judgmental independence, which were crucial factors contributing to enhanced learner autonomy.

"... They should compare the AI's translation with their own draft, then analyze differences in word choice, grammar, and style. Second, I encourage them to check accuracy by reading the original text carefully and asking whether the AI output truly reflects the meaning. For example, if ChatGPT translates an idiom literally, students must decide whether to keep it or adapt it to natural Vietnamese or English usage. Finally, I ask them to revise the AI-generated text to match academic standards, because translation is not only about meaning but also about tone, context, and appropriateness. This way, they learn to use AI critically and responsibly."

5. Discussion

This research paper attempts to seek answers to the following questions:

RQ1: What are the factors influencing English-majored students at HUFLIT to adopt ChatGPT in their translation courses?

Regarding the first research question, the quantitative data obtained from SMART PLS 3 reveals that the greatest contributory factor to the actual adoption of ChatGPT in translation is intrinsic motivation, closely followed by extrinsic motivation, and social influence. This finding is consistent with that of the research undertaken by Ren (2025). Surprisingly, Perceived Usefulness and Perceived Ease Of Use have little or no significant impact on the actual use of ChatGPT in translation, which is contrary to prior studies whose findings have confirmed the positive influence of these factors on behavioral intention and the actual use of a technological tool (Almogren et al., 2024; Alotaibi et al., 2025; Ren, 2025).

It seems that internal factors are the key determinants driving English majors' use of ChatGPT for translation purposes rather than external utility or ease of use in Vietnamese classrooms. This is justifiable since these surveyed participants are GenZ students, or commonly known as the app generation, who have become accustomed to the use of technological devices such as smartphones and laptops. The use of AI-powered tools like ChatGPT to facilitate language learning has become a social norm in their educational environment; thus, the extent of simplicity or usefulness of the tool does not significantly influence the adoption of this AI assistant tool. According to traditional TAM theory, the straightforwardness of technological adoption is proportionate to the degree of using it; nevertheless, this belief seems no longer true in this context. GenZ students are digital natives, so exploiting a piece of technology is relatively easy for them, particularly when they are studying in an environment conducive to technological developments. This leads to an inconsiderable effect of PEOU on the students' use of ChatGPT in translation.

More importantly, intrinsic motivation such as pleasure or interest is the leading factor behind the students' actual employment of ChatGPT. The Vietnamese students are internally motivated to experiment with the tool in translation perhaps because they find it intriguing to explore the tool's capabilities in a demanding subject, or they feel a sense of fulfillment derived from adopting the chatbot since they can accumulate valuable linguistic knowledge and skills through ChatGPT's offer of diverse translation versions with a wide range of word choice and sentence construction. Another significant determinant behind the chatbot use is extrinsic motivation which is often manifested under different forms of outcome gains such as grades

and rewards. Vietnamese education system is exam-oriented and achievement-driven, and so expectations on attaining high academic performance have been deeply ingrained in generations of Vietnamese students' mentality. The use of ChatGPT can not only allow them to relieve the mounting exam pressure, but it can boost their scores or aid them in accomplishing the desired targets as well.

Social influence is also a factor positively driving students to utilize the tool. The role of this factor has been confirmed in the studies conducted by Almogren et al. (2024), Diep & Dang (2025), and Ren (2025). In a collectivist society characterized by a strong bond between people like in Vietnam, teachers and peers are those whose words of encouragement can change a student's behavior. Moreover, teachers are often seen as revered authority figures who can strongly shape students' attitude and actions, while students tend to follow peers' example, especially who are highly competent, to utilize a piece of technology that is likely to boost their scores or academic success whether it is simple to use or not. That is why social influence has made a positive impact on the students' actual adoption of ChatGPT.

RQ2: What skills can be explicitly taught to English-majored students as ChatGPT is formally integrated into translation courses?

With respect to the second question, from the viewpoints of the students, their linguistic competency should be enhanced so that they can confidently handle issues related to the exact usage and contextual appropriateness of lexical units and grammar patterns to convey the original meaning of the language in translations produced by ChatGPT. Apart from this, the students are also expected to receive explicit instructions on how to foster their higher-order thinking skills and post-editing competence in translation courses. These views are supported by lecturers who also stress the significance of lower thinking and translation skills. They maintain that foundational strategies such as recall and application should accompany translation skills such as linguistic ability, cultural and contextual awareness, topical expertise, and techniques like rewording, text simplification, and register use. These findings are in alignment with translation competence model and metacognitive theory. Translation competence, developed by the PACTE group in 2003, is defined as the capability of translating textual information with skills and expertise knowledge (Wilss, 1996, cited in Salamah, 2021) encompassing five elements, namely "language competence, textual competence, subject competence, cultural competence, and transfer competence" (p. 279). Meanwhile, metacognition is an indispensable part of translation competence, in which learners are required to think about the topic in advance as well as reflecting on it during the learning process (Posner et al., 1982, cited in Yanqun, 2015). Yanqun (2015) asserts that this kind of reflection can greatly assist students to achieve academic success as they can look back at their experiences, learning activities, and decisions for in-depth analysis to identify areas for improvement so that they are able to perform better in their future learning paths. In the research context, it is observed that Vietnamese students often struggle with translation due to their limited size of lexeme and grammar knowledge as well as weaknesses in applying diverse skills to translation. The frequent adoption of AI tools can lead to more constant exposure to accurate English; however, they should have their knowledge and translational skills consolidated to double-check the quality of ChatGPT-generated translations, from which they can refine the texts on their own. This matches well with the content of the translation competence model and metacognitive theory discussed above.

RQ3: What kinds of pedagogical instructions or strategies could be incorporated into translation teaching when ChatGPT becomes a formal tool for teaching and learning English?

In terms of the types of pedagogical instructions that can be integrated into translation teaching when ChatGPT becomes a formal tool for language education, the lecturers suggest providing specific, transparent guidance on how to utilize the chatbot appropriately and productively for self-study and in-class practice along with imposing strict constraints on its adoption in tests and forms of assessment. This is justifiable since an increased reliance on AI-powered tool is bound to undermine students' translation competence, and this can be even worse in the context of Vietnam where expectations for outperformance in exams are often appreciated at the expense of critical thinking skills and judgmental capability development. Simultaneously, it is lecturers who should promote students' autonomy and accountability for using ChatGPT, helping them recognize biases and other relevant ethical issues. These concur with that indicated by Almogren et al. (2024) who underscored how ChatGPT could be utilized in a more responsible and ethical manner in educational settings.

When it comes to pedagogical strategies, the lecturers who attended the interview believe that mindset change and role shifting are essential transformations they can make to benefit students' translation learning process. Embracing the utilization of ChatGPT in translation instead of prohibiting it will allow better decision-making skills among students, while lecturers acting as a facilitator or scaffolder can help learners bridge the gaps in their knowledge and skills through guided questions. Adaptations to the lesson content such as interaction enhancement and incorporation of context and meaning-focused activities, as well as those cultivating critical translation and comprehension skills should be made once ChatGPT has been formally integrated into translation courses. The underlying reason for this is translation classes in Vietnam is often teacher-centered, characterized by lecturers transferring knowledge and students passively involved in the lessons. The addition of more interactivity-boosting activities can bring fresh air to the class, changing its focus to a more student-centered one, which in turn encourages learners to take a more active role. Translation is not only mechanically conveying the meaning of the original sentences in the target language, but it also includes analyzing and understanding the sentence meaning and its context in depth, and to achieve this objective, students need to dynamically participate in a variety of tasks requiring critical evaluation, error judgement, reflection, and collaboration. More importantly, students' ability to acquire a language could be greatly enhanced through conversations with their teachers and peers. This finding ties well with sociocultural theory proposed by Vygotsky (1978) and scaffolding technique, which is an integral part of the Zone of Proximal Development model (ZPD). In the words of Vygotsky (1978, cited in Lee, 2015), students' language acquisition process is mainly constructed through their interactions with peers and teachers, and this belief, together with the Zone of Proximal Development model (ZPD), emphasizes the accomplishment of learning under guidance of teachers or through communication of learners with those who are more knowledgeable than them in a specific discipline (Alghamdy, 2023; Muntasir & Akbar, 2023). In the meantime, scaffolding technique revolves around teachers' provision of assistance or pedagogical instruction to students, with the aim of fostering their engagement in learning (Sarmiento-Campos et al., 2022). According to these researchers, scaffolding means more than mere ordinary support, but specific, accurate support given promptly can become a source of motivation leading to an increase in active involvement among learners. Vietnamese students are inclined to be reserved, so constructive support from their teachers or more capable friends are of great benefit to them. Taking on the role of an academic guide or a scaffolder enables teachers to keep track of their students' progress and implement meticulous plans for students' knowledge reinforcement.

Implications for Vietnam's ESL Transition

- **Educators**

Lecturers should add elements of interactivity to lessons to enhance students' engagement during the course. They should switch to the role of a scaffolder or facilitator to maximize students' learning productivity since merely letting them depend on the tool does not guarantee academic progress as they still struggle with a diversity of linguistic aspects that they are not competent in. Group-based projects and student-centered or peer-supported activities could be leveraged to increase the use of ChatGPT in translation due to the strong influence of capable peers on others. However, educators should also design lessons nurturing students' fundamental linguistic knowledge, translation skills, higher-order thinking skills, and post-editing competence so that students have the capability of translating the assigned texts individually with great confidence.

- **Curriculum design**

The research findings imply that curriculum design for translation courses in Vietnamese ESL-oriented context could place an emphasis on enhancing motivational factors among students since these significantly influence the use of ChatGPT in translation. Both intrinsic and extrinsic motivation are closely intertwined and contribute to the degree of actual employment of the chatbot. Curriculum designers, therefore, need to be fully aware of the impact of these factors to incorporate modules encouraging students to use the AI-powered tool with genuine interest or pleasure since stimulation of authentic interest can result in a more active attitude, which entails deeper learning. Making them understand the real value of intrinsic motivation in academic studies could also free them from excessive reliance on the tool for external benefits, thereby gearing them towards viewing translation learning as a worthwhile experience. Employment benefits should be emphasized to serve as external motivation for increasing the adoption of AI tools as well. If the students are well informed about the increasing demand for technological literacy in the labor market, they will be more motivated to utilize AI-powered tools to land a high-paying job.

Separate modules on higher-order thinking skills and post-editing competency should also be integrated into translation courses to provide students with training on how to draft effective prompts and critically evaluate a piece of translation for better error identification and rectification, specifically when it is pertinent to cultural or contextual subtleties. Indeed, when English has become a second language for instruction in tertiary education settings, the capabilities to analyze and judge textual translations produced by AI tools with a critical eye seem to be of paramount importance, which may help differentiate AI-generated versions and those written by humans.

6. Conclusion

The factors influencing English-majors' employment of ChatGPT in translation have been examined using SMART-PLS 3 and structured interviews in this research paper. The study results reveal the varying effects of Perceived Ease Of Use (PEOU), Perceived Usefulness (PU), Intrinsic and Extrinsic Motivation (IM and EM), and Social Influence (SI) on the students' adoption of ChatGPT in translation courses. While IM, EM, and SI positively affect the use of ChatGPT, PU and PEOU are found to exert an insignificant or practically no impact on the students' employment of this AI assistant tool. The findings also suggest that lecturers in charge of translation subjects should transform or adapt their translation pedagogies through a change in mindset and role as well as an incorporation of activities with a greater focus on interaction, meaning, context, critical translation, and comprehension skills. The study also places an emphasis on the equipment of translation, linguistic knowledge, lower-thinking

and higher-order thinking skills, an increased awareness of contextual and cultural subtleties, as well as post-editing competence for English majors. However, lecturers should provide students with appropriate pedagogical guidance concerning the acceptance of ChatGPT adoption in classrooms for linguistic development and guarantee the imposition of restrictions on its use in exam conditions, along with consciousness raising of related ethical issues in employing the chatbot.

Limitations

The study reveals some limitations, with the first being the sample size. With a larger sample size, the obtained results would be more generalizable, thereby increasing the validity and reliability of the findings. Another shortcoming is the employment of structured interviews in the study could limit in-depth insights into the participants' perspectives on the research questions.

Recommendations for further studies

Further studies should consider the influence of a mediating variable on the relationship between the above-mentioned factors, especially Perceived Usefulness, and the actual use of ChatGPT in translation. Within the research context, the correlation between PU and AU is not significant, which contradicts the findings of previous studies.

References

Abdallah, N., Katmah, R., Khalaf, K., & Jelinek, H. F. (2025). Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration. *Social Sciences & Humanities Open, 12,* 101866. https://doi.org/10.1016/j.ssaho.2025.101866

Akil, M. (2011). The Quality Of Indonesian - English Translation By English Department Students Of Higher Learning Institutions In Makassar. *Linguistika: Buletin Ilmiah Program Magister Linguistik Universitas Udayana, 18*(1), 1-14.

Albalawi, A. F., & Abdul-Ghafour, A.-Q. K. (2026). An investigation into human vs AI English translations of Qur'anic euphemisms. *Social Sciences & Humanities Open, 13*, 102249. https://doi.org/10.1016/j.ssaho.2025.102249

Alghamdy, R. Z. (2023). English Teachers' Practice of Classroom Discourse in Light of Zone of Proximal Development Theory and Scaffolding Techniques. *Journal of Language Teaching and Research, 15*(1), 46-54. https://doi.org/10.17507/jltr.1501.06

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on Learning Motivation: Teachers and Students' Voices. *Journal of English Studies in Arabia Felix, 2*(1), 41-49. doi: 10.56540/jesaf.v2i1.51

Aljabr FS, Al-Ahdal AAMH. (2024). Ethical and pedagogical implications of AI in language education: an empirical study at ha'il university. *Acta Psychol (Amst)* (PubMed / Europe PMC), 251, 104605. https://doi.org/10.1016/j.actpsy.2024.104605

Al-kfairy, M. (2024). Factors Impacting the Adoption and Acceptance of ChatGPT in Educational Settings: A Narrative Review of Empirical Studies. *Applied System Innovation, 7*(110). https://doi.org/10.3390/asi7060110

Almansour, N. (2024). Investigating Factors Influencing EFL Learners' Behavioral Intentions to Adopt ChatGPT for Language Learning. *Journal of Language Teaching and Research, 15*(6), 1910-1918. https://doi.org/10.17507/jltr.1506.16

Almogren, A. S., Al-Rahmi, W. M., & Dahri, N. A. (2024). Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon*, *10*(11), e31887. https://doi.org/10.1016/j.heliyon.2024.e31887

Alotaibi, H. M., Sonbul, S. S., & El-Dakhs, D. A. (2025). Factors influencing the acceptance and use of ChatGPT among English as a foreign language learners in Saudi Arabia. *Humanities & Social Sciences Communications, 12*(1). https://doi.org/10.1057/s41599-025-04945-2

Amin, M. Y. (2023). AI and Chat GPT in Language Teaching: Enhancing EFL Classroom Support and Transforming Assessment Techniques. *International Journal of Higher Education, 4*(4), 1-15. https://doi.org/10.33422/ijhep.v4i4.554

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Salhi, N. R. A. (2023). Exploring English language learning via Chabot: A case study from a self determination theory perspective. *Computers and Education: Artificial Intelligence*, *5*, 100148. https://doi.org/10.1016/j.caeai.2023.100148

Badah, A. M., Khalaf, C. N., Dwaikat, F. J., & AlQbailat, N. (2026). The usage of artificial intelligence in legal translation: bridging the gap between law and language. *Ampersand*, *16*, 100248. https://doi.org/10.1016/j.amper.2025.100248

Daskalakis, S., & Mantas, J. (2008). Evaluating the impact of a service-oriented framework for healthcare interoperability. *Studies in Health Technology and Informatics*, *136*, 285.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*, New York: Plenum.

Diep, M. C., & Dang, M. H. (2025). Exploring factors influencing the adoption of chatgpt as a supportive tool in EFL lesson creation. *AsiaCALL Online Journal*, *16*(1), 111-139. https://doi.org/10.54855/acoj.251616

Dong, L., Liu, M., & Yang, F. (2022). The Relationship Between Foreign Language Classroom Anxiety, Enjoyment, and Expectancy-Value Motivation and Their Predictive Effects on Chinese High School Students' Self-Rated Foreign Language Proficiency. *Frontiers in Psychology, 13*, 860603. doi: 10.3389/fpsyg.2022.860603

Fatoni, A. U. (2022). An analysis of online learning media in english class during covid-19 pandemic. doi:10.2991/assehr.k.220407.122

Ferreiro-Santamaria, G. (2024). Exploring the Role of ChatGPT in English Teaching Within Higher Education Settings. *International Journal of Trends and Developments in Education, 4*(1), 44-58. doi:10.5281/zenodo.11204781

Fitria, T. N. (2018). Translation Techniques Found In English To Indonesian Abstract Translation Of Journal Edunomika 2018. *Elite: English and Literature Journal*, *5*(2), 145-160. https://doi.org/10.24252/elite.v5i2a4

Ghimire, P. R., Neupane, D. B., & Dahal, N. (2024). Generative AI and AI Tools in English Language Teaching and Learning: : an exploratory research. *English Language Teaching Perspectives*, *9*(1-2), 30-40. https://doi.org/10.3126/eltp.v9i1-2.68716

Gutiérrez, L. M. (2023). Artificial Intelligence in Language Education: Navigating the Potential and Challenges of Chatbots and NLP. *Research Studies in English Language Teaching and Learning (RSELTL), 1*(3), 180- 191. https://doi.org/10.62583/rseltl.v1i3.44

Hadalgekar, M. S., & Desai, D. N. (2025). Role of Perceived Ease of Use and Perceived Usefulness in Adoption of Mobile Computing Technology. *International Journal of Social Science and Human Research, 8*(6), 4458-4462. doi: 10.47191/ijsshr/v8-i6-58

Hair, J. F., Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling* (2nd Ed.). Thousand Oaks, CA: Sage

Hair, J.F., Hult, G.T.M., Ringle, C.M. & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling* (PLS-SEM). Thousand Oaks, CA: Sage.

Harryanto., Muchran, M., & Ahmar, A. S. (2018). Application of TAM model to the use of information technology. *International Journal of Engineering & Technology*, *7*(2.9), 37–40. https://api.semanticscholar.org/CorpusID:59523621

Henseler, J., Ringle, C. M., & Sarstedt, M. (2012). Using partial least squares path modeling in international advertising research: Basic concepts and recent issues. In S. Okazaki (Ed.), *Handbook of research on international advertising* (pp. 252–276). Edward Elgar Publishing.

Henseler, J., & Sarstedt, M. (2012). Goodness-of-fit indices for partial least squares path modeling. *Computational statistics*, *28*(2), 565-580. doi: 10.1007/s00180-012-0317-1

Höck, M., & Ringle, C. M. (2010). Local strategic networks in the software industry: an empirical analysis of the value continuum. *International Journal of Knowledge Management Studies*, *4*(2), 132-151.

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification, *Psychological Methods, 3*(4), 424-453. https://doi.org/10.1037/1082-989X.3.4.424

Ismail, H., Syahruzah, J. K., & Basuki, B. (2017). Improving The Students' Reading Skill Through Translation Method. *Journal of English Education*, *2*(2), 124-131. https://doi.org/10.31327/jee.v2i2.405

Jeno, L. M., Raaheim, A., Kristensen, S. M., Kristensen, K. D., Hole, T. N., Haugland, M. J., & Mæland, S. (2017). The relative effect of team-based learning on motivation and learning: A self-determination theory perspective. *CBE—Life Sciences Education*, *16*(4), ar59. https://doi.org/10.1187/cbe.17-03-0055

Jukiewicz, M., & Wyrwa, M. (2026). Can ChatGPT Replace the Teacher in Assessment? A Review of Research on the Use of Large Language Models in Grading and Providing Feedback. *Appl. Sci., 16*(680). https://doi.org/10.3390/app16020680

Kartini, M. M., Santosa, M. H., & Indrayani, L. (2025). Students' Acceptance and Enjoyment of ChatGPT in English Language Learning. *Journal of Educational Study (JoES), 5*(2), 171-186. https://doi.org/10.36663/joes.v5i2.1100

Kiet, A. (2025, November 4). *English to be Vietnam's second language by 2035*. *Hanoi Times*. https://hanoitimes.vn/english-to-be-vietnam-s-second-language-by-2035.893519.html

Lai, P. (2017). The Literature Review Of Technology Adoption Models And Theories For The Novelty Technology. *Journal of Information Systems and Technology Management, 14*(1), 21-38. doi: 10.4301/S1807-17752017000100002

Lee, M. S. (2015). Implementing the Sociocultural Theory While Teaching ESL. *SPACE: Student Perspectives About Civic Engagement, 1*(1), 28-35. Retrieved from https://digitalcommons.nl.edu/space/vol1/iss1/6

Li, M., Enkhtur, A., Cheng, F., & Yamamoto, B. A. (2024). Ethical implications of ChatGPT in higher education: A scoping review. *Journal of Interdisciplinary Studies in Education*, *13*(1), 55-69. doi: 10.32674/jise.v1

Li, X., Lin, B., Zhao, W., Jin, Q., Lin, T., & Li, Z. (2025). Research and Practice of ChatGPT-Assisted Instructional Design for University Teachers. *International Journal of New Developments in Education*, *7*(3), 18-24. doi: 10.25236/IJNDE.2025.070304

Manapbayeva, Z., Zaurbekova, G., Ayazbekova, K., Kazezova, A., & Pirmanova, K. (2024). AI in Literary Translation: ChatGPT-4 vs. Professional Human Translation of Abai's Poem 'Spring'. *Procedia Computer Science 251*, 526–531. doi:10.1016/j.procs.2024.11.143

Mohsan, M., & Durr-e-Nayab, D. (2024). Estimating and Comparing Translation Skills: A Comparative Study of ChatGPT and Human Translation. *Journal of Development and Social Sciences, 5*(3), 75-86. https://doi.org/10.47205/jdss.2024(5-III)08

Moqbel, M. S., & Al-Kadi, A. M. (2023). Foreign Language Learning Assessment in the Age of ChatGPT: A Theoretical Account. *Journal of English Studies in Arabia Felix, 2*(1), 71–84. doi: 10.56540/jesaf.v2i1.62

Muntasir, M., & Akbar, I. (2023). Revisiting the Significance of ZDP and Scaffolding in English Language Teaching. *Journal of English Language Teaching Linguistics and Literature, 3*(1), 40-45. https://doi.org/10.47766/jetlee.v3i1.1276

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory. (3rd ed.). McGraw-Hill.

Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education, 7*(3), 135-153. https://doi.org/10.2478/jolace-2019-0025

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: towards a theory of conceptual change. *Science Education, 66*, 211-227. https://doi.org/10.1002/sce.3730660207

Qu, K., & Wu, X. (2024). ChatGPT as a CALL tool in language education: A study of hedonic motivation adoption models in English learning environments. *Education and Information Technologies, 29*, 19471–19503. https://doi.org/10.1007/s10639-024-12598-y

Ren, X. (2024). Modeling college EFL teachers' intentions to conduct academic research: Integrating theory of planned behavior with self-determination theory. *PLoS ONE, 19*(8), 1-18. https://doi.org/10.1371/journal.pone.0307704

Ren, X. (2025). We want but we can't: measuring EFL translation majors' intention to use ChatGPT in their translation practice. *Humanities And Social Sciences Communications, 12*(256), 1-11. https://doi.org/10.1057/s41599-025-04604-6

Ringle, C.M., Sarstedt, M., Schlittgen, R., & Taylor, C.R. (2013), 'PLS path modeling and evolutionary segmentation', *Journal of Business Research*, *66*(9), 1318-1324. https://doi.org/10.1016/j.jbusres.2012.02.031

Rousan, R. A., Jaradat, R., & Malkawi, M. (2025). ChatGPT translation vs. human translation: an examination of a literary text. *Cogent Social Sciences*, *11*(1), 1-21. https://doi.org/10.1080/23311886.2025.2472916

Roza, V., & Zulhirawati. (2023). Higher Students' Perception of Using Chat GPT in Translating English Texts. *Bukittinggi International Conference on Education*, *1*, pp. 64-73. https://dx.doi.org/10.30983/biced.v1i1.278

Rubić, I., & Matijević, D. (2019). Autonomy and Motivation: A Self-Determination Theory Perspective on ESP Motivation . *The Journal Of Teaching English For Specific And Academic Purposes, 7*(2), 147-158. https://doi.org/10.22190/JTESAP1902147R

Salamah, D. (2021). Translation Competence and Translator Training: A Review. *International Journal of Linguistics, Literature and Translation, 4*(3), 276-291. doi:10.32996/ijllt.2021.4.3.29

Sarmiento-Campos, N.-V., Lázaro-Guillermo, J. C., Silvera-Alarcón, E.-N., Cuellar-Quispe, S., Huamán-Romaní, Y.-L., Apaza, O. A., & Sorkheh, A. (2022). A Look at Vygotsky's Sociocultural Theory (SCT): The Effectiveness of Scaffolding Method on EFL Learners' Speaking Achievement. *Education Research International*, 1-12. https://doi.org/10.1155/2022/3514892

Semerikov, S. O., Striuk, A. M., & Shalatska, H. M. (2021). AI-assisted language education: critical review. *Educational Dimension, 4*, 1-7.

Sharp, J. H. (2007). Development, Extension, and Application: A Review of the Technology Acceptance Model. *Information Systems Education Journal, 5*(9). https://isedj.org/5/9/

Surya Bahadur G. C., Bhandari, P., Gurung S. K., Srivastava E., Ojha, D., Dhungana, B. R. (2024). Examining the role of social influence, learning value and habit on students' intention to use ChatGPT: the moderating effect of information accuracy in the UTAUT2 model. *Cogent Education*, *11*(1), 1-24. https://doi.org/10.1080/2331186X.2024.2403287

Tafazoli, D. (2024). Exploring the potential of generative AI in democratizing English language education. *Computers and Education: Artificial Intelligence, 7*, 100275. https://doi.org/10.1016/j.caeai.2024.100275

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2016). Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead. *Journal of the Association for Information Systems, 17*(5). https://doi.org/10.17705/1jais.00428

Vietnam's smart shift in education. (2025, October Monday). Retrieved from Tuổi trẻ news: https://news.tuoitre.vn/vietnams-smart-shift-in-education-103251020143536157.htm

Vinola, V., Supriusman, & Syarfi, M. (2025). Perceptions of Students on the Use of ChatGPT as a Translation Tool. *Lectura: Jurnal Pendidikan, 16*(2), 394-405. https://doi.org/10.31849/283kr275

Vygotsky, L. S. (1978). Interaction between learning and development. In *Mind in society: The development of higher psychological processes* (pp. 79–91). Harvard University Press.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology, 25*(1), 68–81. doi: 10.1006/ceps.1999.1015

Wilss, W. (1996). Knowledge and skills in translator behavior. John Benjamins Publishing Company. https://doi.org/10.1075/btl.15

Yanqun, Z. (2015). The Concept and Instruction of Metacognition in Translation Competence Development. *International Forum of Teaching and Studies, 11*(1-2), 69-78.

Yavuz, F., Çelik, Ö., & Çelik, G. Y. (2024). Utilizing large language models for EFL essay grading: an examination of reliability and validity in rubric-based assessments. *British Journal of Educational Technology*, *56*(1), 150-166. https://doi.org/10.1111/bjet.13494

Zhu, M., & Wang, C. (2025). A systematic review of research on AI in language education: Current status and future implications. *Language Learning & Technology, 29*(1), 1-29. Retrieved from https://hdl.handle.net/10125/73606

Bionote

Huong Thi Thu Le is currently an experienced lecturer working for the Faculty of Foreign Languages, HUFLIT, Vietnam with more than twelve years of teaching experience. She earned a bachelor's degree from Ho Chi Minh University of Education and holds a master's degree in TESOL from Victoria University. Her interests are primarily novel methodologies to enhance students' learning and achievement, AI tools, and how these can make radical improvements in learners.

Appendices

Appendix A: Questionnaire

EXAMINING THE FACTORS AFFECTING THE EMPLOYMENT OF CHATGPT IN TRANSLATION AMONG ENGLISH-MAJORS AT HUFLIT AS VIETNAMESE EDUCATION TRANSITIONS FROM EFL TO ESL ORIENTATION

Dear students, I am conducting a study to investigate the factors behind the adoption of ChatGPT in translation among English majors at HUFLIT as Vietnam shifts from an EFL to ESL-focused education. I would love you to spend some of your time completing this survey, which should last approximately 10 minutes. I am sure that your responses are kept completely confidential. Thank you very much for your participation.

SECTION 1: BACKGROUND INFORMATION

Email Address

Which gender are you?

- o Male
- o Female
- o Prefer not to say

Which major are you in?

- o Translation - Interpretation
- o Business English
- o English Pedagogy
- o English for Office Administration
- o English for Logistics
- o English - Chinese Bilingualism

Have you ever used ChatGPT for translation exercises?

- o Yes (Move to the next question)
- o No (You can stop here. Thank you for your participation)

How often do you use ChatGPT in translation subjects?

- o Always
- o Usually
- o Sometimes
- o Seldom

SECTION 2: PERCEIVED USEFULNESS OF CHATGPT

Rate the following statements on the perceived usefulness of ChatGPT on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5

1	ChatGPT enhances my performance in translation courses.					
2	ChatGPT's customized feedback helps me improve the accuracy of translated texts.					
3	ChatGPT gives me multiple options of vocabulary, grammar, and idea expressions, which helps improve the quality of my translations.					
4	ChatGPT assists me to complete translation exercises more effectively.					

SECTION 3: PERCEIVED EASE OF USE

Rate the following statements on the perceived ease of use of ChatGPT on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5
1	ChatGPT's interface is user-friendly, making it easy for me to use it for translation tasks.					
2	Responses from ChatGPT to my requests are clear and understandable.					
3	Giving prompts to ChatGPT to translate texts is simple.					
4	I find it simple to learn how to use ChatGPT for translation exercises.					

SECTION 4: MOTIVATION BEHIND THE ADOPTION OF CHATGPT

INTRINSIC MOTIVATION

Rate the following statements on the intrinsic motivation behind the employment of ChatGPT on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5
1	I find it interesting to use ChatGPT for translation assignments.					
2	Adopting ChatGPT enables me to take greater control of my translation learning.					
3	I enjoy using ChatGPT to generate different translation versions.					
4	I am curious to find out how ChatGPT can assist me with translation work.					

SECTION 5: MOTIVATION BEHIND THE ADOPTION OF CHATGPT

EXTRINSIC MOTIVATION

Rate the following statements on the extrinsic motivation behind the employment of ChatGPT on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5
1	I use ChatGPT to achieve better results in translation courses.					
2	Adopting ChatGPT helps me reduce workload pressure in translation courses.					
3	I can better cope with assessment requirements thanks to the use of ChatGPT.					

SECTION 6: SOCIAL INFLUENCE

Rate the following statements on the social influence of using ChatGPT in translation on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5
1	People who are important to me think I should employ ChatGPT for translation.					
2	My lecturer encourages me to use ChatGPT in translation.					
3	I feel I am influenced to adopt ChatGPT for translation assignments by my peers.					

SECTION 7: CHATGPT EMPLOYMENT IN TRANSLATION

Rate the following statements on ChatGPT employment in translation on a scale of 5 (1 = Completely Disagree; 2 = Disagree; 3 = Neither Agree nor Disagree; 4 = Agree; 5 = Completely Agree)

No.	Statements	1	2	3	4	5
1	I frequently use ChatGPT when handling translation assignments.					
2	I rely on ChatGPT as a tool for supporting my translation work.					

Appendix B: Interview Questions

FOR STUDENTS

1. What kind of translation assignments do you often use ChatGPT for?
2. What skills do you think are required among students of English when ChatGPT is formally used in translation courses? Can you give me an example?

FOR LECTURERS

1. What types of instruction do you often give to your students who intend to use ChatGPT in their translation exercises?
2. How has the students' adoption of ChatGPT changed the way you teach translation subjects?
3. What kinds of activities or strategies do you think can be incorporated into translation teaching when ChatGPT becomes a formal tool for teaching and learning?
4. What skills can be explicitly taught to students as ChatGPT is formally integrated into translation courses?

#48 Investigating Students' Perspectives and Engagement Levels with AI Tools in Second Language Acquisition (SLA) Courses

Lan DoThi

Abstract

In the fast-growing digital era, the integration of Artificial Intelligence (AI) tools has become indispensable in higher education, particularly in language studies where these technologies offer personalized linguistic support and cognitive scaffolding. Despite this prevalence, empirical evidence regarding student engagement in specialized subjects like Second Language Acquisition (SLA) remains limited. This study employs a quantitative and qualitative approach, surveying 125 fisrt-year students at HUFLIT upon their course completion to examine their perspectives and usage patterns of AI applications. Data analysis, conducted via Microsoft Excel, reveals a profound correlation between student learning habits and digital tool engagement, indicating that learners increasingly rely on AI for conceptual clarity and task execution. The findings suggest that AI tools are no longer optional, but they are seen as central to the modern learning experience. Consequently, it is recommended that educators proactively update their pedagogical strategies to align with these digital trends. If properly adopted, embracing AI-driven methodologies will support instructional efficacy in teaching and foster a more responsive, tech-integrated environment for second language learners.

Key words*: Second Language Acquisition (SLA), strategies, applications, AI-driven, efficacy*

Abbreviations:

SLA: Second Language Acquisition

AI: Artificial Inteligence

Q: Question

L1: First Language

L2: Second Language

I hereby formally confirm my commitment to serve as a speaker at the HUFLIT INTESOL INTERNATIONAL CONFERENCE 2026 conference. I acknowledge that failing to uphold this commitment without valid prior notice will result in my exclusion from participation in the next two consecutive HUFLIT INTESOL conferences.

1. Introduction

1.1. Rationale for choosing the research topic

With the trend of integration and globalization, language learning has become a necessity today. The languages we teach will face a major challenge: the emergence of artificial intelligence (AI). In the digital age AI is a transformative force, driving efficiency and innovation across all sectors. In other words, AI is no longer a luxury but a necessity; it provides personalized paths, instant feedback, and immersive environments that traditional teaching methods often lack. By leveraging AI tools, learners can achieve fluency faster, bridging cultural gaps and mastering Second Language Acquisition (SLA) with unprecedented precision

and scale. However, certain difficulty in learning SLA as a subject for the first-year students at Huflit may raise a question.

While previous studies have extensively discussed the technical capabilities of AI, there remains a critical research gap regarding the human element. This study, by investigating students' perspectives and engagement levels with AI tools in Second Language Acquisition (SLA) courses, seeks to uncover the subjective experiences that determine how effective these tools are in the SLA classroom.

1.2. Research Objectives

This research aims to comprehensively evaluate the integration of Artificial Intelligence (AI) among first-year students within the context of Second Language Acquisition (SLA). Specifically, the study seeks to investigate students' attitudes toward AI-driven tools, examining their perceptions of usefulness and ease of use. Furthermore, it analyzes the specific behaviors and engagement patterns exhibited by learners during the language learning process. Ultimately, the research assesses the effectiveness of these technologies in improving linguistic proficiency and learning outcomes. By synthesizing these three factors, the study provides critical insights into how AI can be optimized to enhance foundational language education in the digital era.

1.3. Research questions

To achieve the research objectives, the following three core questions have been formulated:

(1) What are the perceptions and attitudes of first-year students toward the integration of AI in SLA?

(2) What are the predominant AI tools and usage patterns exhibited by these students in their language learning routines?

(3) To what extent do first-year students perceive AI as effective in enhancing their linguistic proficiency and overall learning outcomes?

1.4. Scope of Research

The research subjects are first-year students at Huflit University. They are first-year students who have just finished the course titled "An Introduction to SLA". Each class has approximately 70 students. However, 125 was the total number of respondents in the two classes surveyed at the end of the course.

2. Literature Review

2.1. SLA and the significance of technology

2.1.1. Definition of SLA

According to Benati & Angelovska (2016), SLA is the study of how learners create a new language system despite an often-limited exposure to the second language/foreign language learning contexts. In a general sense, a second language (L2) refers to a language that is acquired after the first language (L1) has been established in early childhood.

SLA addresses two fundamental questions: (1) how learners come to internalize the linguistic system of another language; (2) how learners make use of that linguistic system during comprehension and speech production (Benati & Angelovska, 2016).

The following are distinctions commonly made in literature:

"A second language" is typically an official or societally dominant language needed for education, employment, and other basic purposes. It is often acquired by minority group members or immigrants who speak another language natively. In this more restricted sense, the term is contrasted with other terms. However, the term "a foreign language" is one not widely used in the learners' immediate social context which might be used for future travel or other cross-cultural communication situations or studied as a curricular requirement or elective in school, but with no immediate or necessary practical application (Saville-Troike & Barto, 2017).

2.1.2. SLA learners' purposes

There are six broad groups of L2 users: People who are part of multilingual communities; People regaining their cultural heritage, short-term visitors to another country; People using an L2 with partners, friends or children; People using an L2 internationally for specific functions and People using an L2 internationally for a wide range of functions (Cook & Singleton, 2014).

The learner's purpose in acquiring the L2 in some cases is not related to long-term integration only, but has more practical, short-term goals - such as meeting a requirement for school or university matriculation or graduation, applying for a job, or gaining access to material in the second language (Cook & Singleton, 2014).

The use of a language - whether first or second - has as its goal the communication of ideas, maintenance of social relations, and creation of discourse, all of which require the development of several core competencies. These include input competencies, such as listening and reading, and output competencies, such as speaking and writing, and entail the process of receiving (input), attending to (interaction), and assigning meaning (output) to verbal (aural or written) and/or visual stimuli. They also include competencies in communicating in the target language, which includes an understanding of the cultural context of such communication (Jones, 2014).

2.1.3. The Impact of Technology as Mediator of SLA Processes

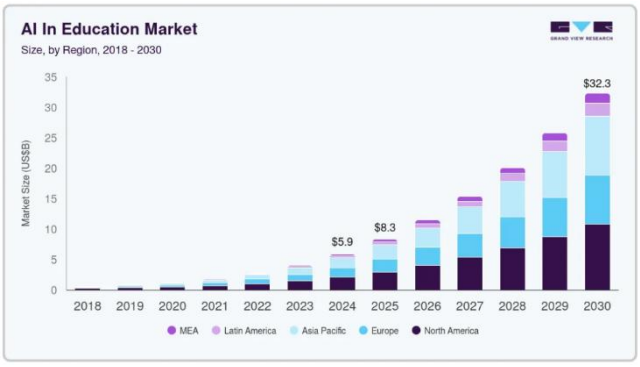
In teaching, especially language teaching, adopting technology properly as a language input is expected by both intructors and learners. The enhancement of learning outcomes is practically possible since "technology allows for exposure to authentic target language input, in the form of multimodal websites that contain text-based, audio, and video information about both language and culture. In addition, many different types of Computer-Mediated Communication, often classified as Web 1.0 and Web 2.0 technologies (e.g., email, forums, chats, blogs, wikis, podcasts, videos, social networking sites, virtual worlds, massively multiplayer online games) enable various kinds of communication, interaction, and collaboration" (Chun, 2016).

2.2. AI when used in Education

2.2.1. Advantages of AI tools in education

It is undeniable that AI tools have proven their advantages, not only from the teachers' perspective but also from the learners'. Tseng & Warschauer (2023) claimed: "Even if we could ban AI from education, we shouldn't. ChatGPT and similar tools offer powerful affordances for second language learners that should not be discounted".

Looking at this chart, what we can see is the irrevesable trend. The global AI education market is predicted to exceed $32.3 billion by 2030. What we may infer is that these are not simply forecast figures, but rather an assessment of a global development trend.



Source: Grand View Research, AI In Education Market Sise & Share

The question is whether we may use AI as effectively as we wish, or will AI somehow influence our teaching, change learners' perspectives, and force us to adapt our teaching methods to keep pace with the rapid development of AI and its applications.

In terms of accessibility and functionality, there are clear advantages that these tools can offer. According to AIT (2023), if we apply AI Chatbot in education, AI Chatbot will act as an assistant, helping teachers with many time-consuming tasks such as imparting knowledge, grading, guiding students to solve some basic exercises, storing information about scores, and monitoring students' learning progress. ChatGPT's writing style is concise, smooth and very natural. The answers are written neatly and concisely, without spelling mistakes. Learners must be smart to rely on ChatGPT to exploit the Bot to bring positive benefits to themselves (An, 2023).

ChatGPT can affect several aspects of education, including writing, instruction methods and teaching pedagogy (Javaid et al., 2023). Aktay et al. (2023) believe that ChatGPT is its ability to respond rapidly. Education and AI are critical topics of conversations about the future of our society.

According to Rahman and Watanobe (2023), ChatGPT can help students solve complex problems. One of the most powerful features of ChatGPT is its human-like understanding and response. Learners can now ask ChatGPT questions in the same way as they would ask their tutor. From elementary school to higher education and professional development, it can be used at all levels of education. However, ChatGPT's extraordinary ability to perform complex tasks in the education field has sparked controversy while some are considering ChatGPT and AI to be the future of teaching and learning, as well as educational research (Baidoo-Anu & Ansah, 2023). Rasul et al., (2023) claimed that as ChatGPT continues to improve and evolve, it is likely to become an increasingly important tool for researchers and scholars and promises to be a potential application for integration into education in the coming years.

Karanjakwut and Charunsri (2025) pointed out that before the introduction of ChatGPT, traditional brainstorming was still a widely used technique in the writing process, which involved group discussions or mind mapping and listing ideas to generate many different thoughts and solutions. However, the method limits the search space, decreasing the diversity of ideas (Oanh, 2023). Thus, ChatGPT has become a potential tool to facilitate brainstorming and identification of relevant reference materials during the research process (Mutiara et al.,

2024). Furthermore, ChatGPT can help suggest article titles, expand or shorten articles, and even suggest creative ideas to improve the quality of articles (Zebua & Katemba, 2024).

Of course, AI tools may help thanks to Computer Mediated Communication (CMC). Language learners can communicate directly, inexpensively, and conveniently with other learners or speakers of the target language 24 hours a day, from school, work, or home. This communication can be asynchronous (not simultaneous) through tools such as electronic mail (email), which allows each participant to compose messages at their time and pace, or in can be synchronous ("real time"), which allow people all around the world to have a simultaneous conversation by typing at their keyboards. It also allows not only one-to-one communication, but also one-to-many, allowing a teacher or student to share a message with a small group, the whole class, a partner class, or an international discussion list of hundreds or thousands of people (Warschauer, 1996).

Advocates of the use of various kinds of technology-enhanced instruction have pointed to studies which suggest that students enjoy learning under such conditions, and find it less pressuring than traditional teaching, leading to an improved attitude towards the language-learning task. A fair amount of research has shown that a good experience of and/or success in learning a particular language can bring about favourable attitudes (Cook & Singleton, 2014).

2.2.2. Disadvantages and problems with AI Tools

According to Dziak (2024), some people believe that technology takes jobs away from human workers. Other people have resisted new technology because of its high cost, which may represent a major and risky investment for individuals and small businesses. Still, other people have avoided technology because it can be difficult to implement and learn how to use properly.

Chatbots are synchronous tools that support individual learning, however, pointed out that the limitation of technology itself (e.g, low accuracy of chatbots' responses) may diminish the degree to which chatbots can help improve students' performance. (Huang, Hew, & Fryer, 2022).

ChatGPT's feedback accuracy was found to be low, with wording that was often generic, abstract, and thus lacking concrete guidance for improvement. The authors suggested ChatGPT's analyses gave too much weight to the use of transitional words without being able to distinguish between effective and ineffective use of such words (Godwin-Jones, O'Neill & Ranalli, 2024).

It is precisely the abundant availability of not only information, but also misinformation and disinformation, which, if included in the data corpus that ChatGPT has been trained on, may lead to false information in its output. Thus, the possibility that ChatGPT can provide false information will require educators to include a strong focus on critical evaluative skills (Kostka & Toncelli, 2023).

Some uses of AI products in SLA have generally been seen positively (such as Automatic Writing Evaluation), the use of others has been controversial, with some language teachers, for example, attempting to ban student use of machine translation (Godwin-Jones, O'Neill & Ranalli, 2024).

Let's talk about contextual adaptation. In terms of contextual and cultural nuances. Culture provides the necessary context for accurate translation. It enables professional language translators to make informed decisions regarding word choices, tone, and style that align with the cultural expectations of the target audience (Ccelingo, 2023). However, students

do not always remember this element. In recent years, advancements in machine translation technology, such as Google Translate, and AI have made significant strides in facilitating cross-linguistic communication. However, these tools often struggle to capture the emotional subtleties and nuances inherent in human language (Najimova, 2024).

3.1. Methodology

3.1. Research Design

Since the purpose of the study was to explore the students' perspectives and engagement levels with AI Tools in Second Language Acquisition (SLA) Courses and to get more recommendations accordingly, this study employs a questionnaire as a cross-sectional survey to provide a comprehensive understanding of student engagement with AI in the subject SLA. The questionnaire was used to quantify students' attitudes and perceptions of AI usefulness, specific engagement behaviors (Quantitative) and to learn more about students' recommendations (Qualitative). This approach hopefully ensures an expectedly good assessment of how AI integration affects linguistic proficiency and overall learning outcomes.

3.2. Description of Subjects

The participants of this research are 125 first-year students from the Faculty of Foreign Languages at HUFLIT University. These participants have specifically completed their Second Language Acquisition (SLA) courses, ensuring they possess the necessary direct experience with AI tool integration to provide informed perspectives on engagement and learning outcomes.

3.3. Instrument

It is a questionnaire form that consisted of 14 questions where likert-scales were used. 12 selected-response questions (close-ended questions) served as a quantitative measure, and 02 constructed-response questions (open-ended questions - Q6 and Q14) served as a qualitative measure to learn more about recommendations made by the students as survey participants.

3.4. Procedure

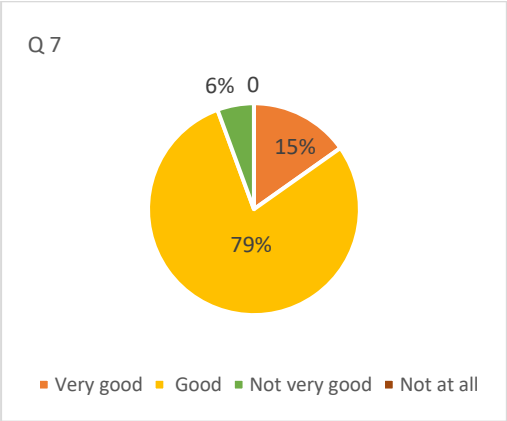
First, a questionnaire was cteated. Next, the forms were distributed to the two classes at the end of the course, namely, the last session. The students, as the survey participants, were allowed to spend 15 minutes completing the forms without providing their personal information. This is to ensure that they may feel free to contribute their points into the topic. The feedback of the respondents was colledted accordingly. After the data collection, the statistics were demonstrated and processed through MS Excel for information management and interpretation.

4. Findings

4.1. Digital Literacy and AI Readiness

This study addresses three primary research questions aligned with the stated objectives. However, the responses to Q7 "How would you rate your proficiency in using AI tools?" revealed the proficiency of the students, which was the baseline for the three others.

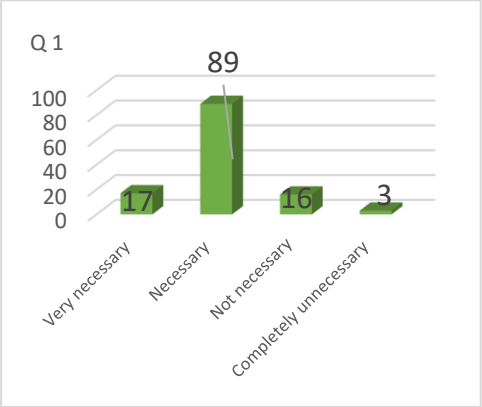
Fig.1 - Proficiency in using AI tools/ Apps



A significant majority of participants (94%) reported high proficiency in AI usage, with 15% rating themselves as "Very Good" and 79% as "Good". That claim shows the "HUFLIT English majors" in the study appear to be "high perceived digital self-efficacy" who are highly confident in their tech skills.

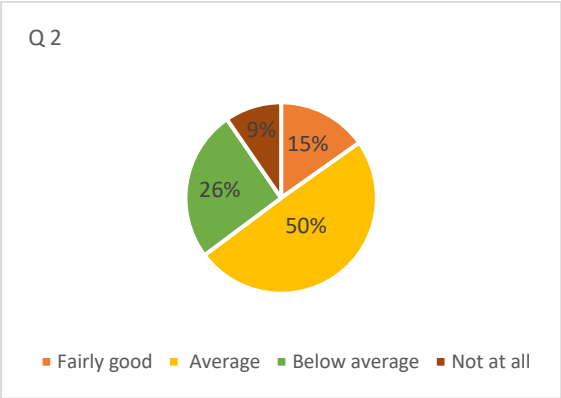
4.2. Perceptions of SLA & AI Necessity (Q1, Q2, Q13)

Fig.2 - The necessity of SLA



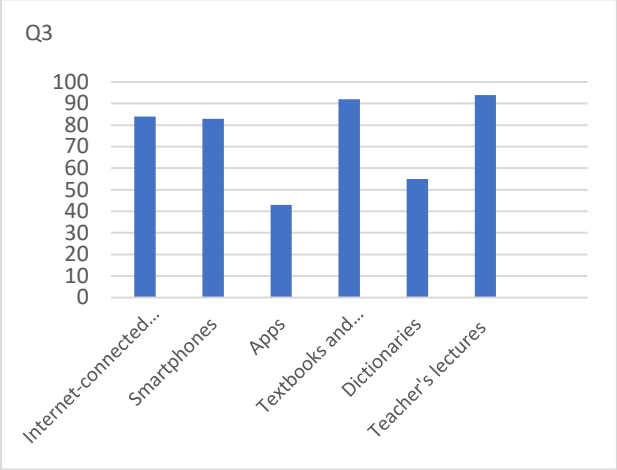
A majority (84%) confirmed the necessity of the subject SLA, with 13% rating as "Very necessary" and 71% as "Necessary". 15% thought that SLA is not. (16 responses indicating "not necessary" and 03 claiming "completely unnecessary")

Fig.3 - The level of understanding of the core knowlegde in SLA.



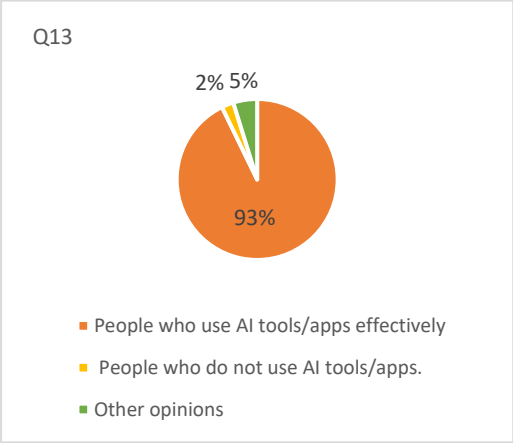
65% confirmed their understanding of the lessons (with 15% confirming "fairly good" and 50% as "average"). 35% of respondents said their understanding is not as good as expected (with 26% under average level and 09% with no understanding at all).

Fig.4 - Factors helping students' motivation in SLA



84 responses said that one of the factors helping students'motivation is using Internet-connected devices (accounting for 67,2%). Similarly, 83 respondents calimed that smartphones are another helping factor (accounting for 66,4%). However, the number indicating the teachers' lectures as a helping factor is 94 (accounting for 75,2%).

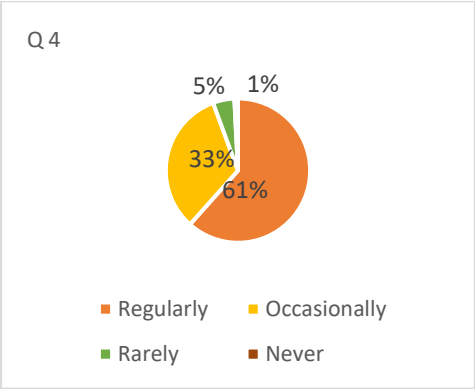
Fig.5 - Perceived learning advantages of AI tool users



93% believed that with AI knowledge, learners may have an advantage in leraning. 07 % did not think so. While 93% students believe AI users have an advantage, only 84% see the subject itself as "Very Necessary" - suggesting a gap between tool utility and subject interest.

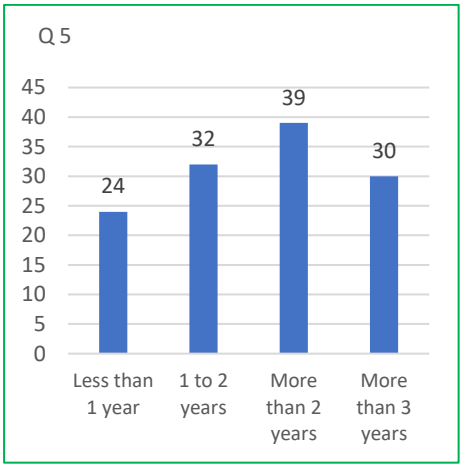
4.3. Patterns of AI Integration (Q4, Q5, Q6, Q8, Q9)

Fig.6 - Frequency of consulting AI Tools/ AI Apps for SLA



61% said that they regularly consulted AI Tools for the subject. 33 % - occasionally and 05% - rarely referring to AI.

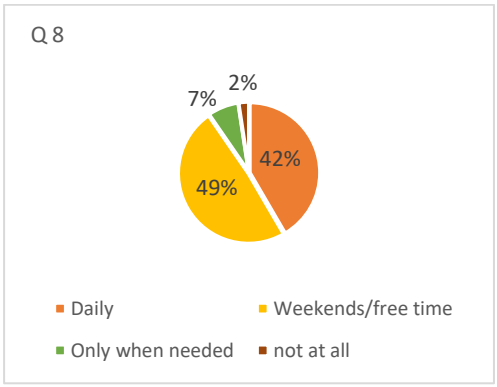
Fig.7 - Respondents' history with AI in education



19% stated that they have known GenAI for less than 1 year. 81% claimed that they have known AI tools for more than one year.

Q6 - When students were asked to list the names of the AI tools that they use for the SLA subject, they affirmed the "Big Three" tools (ChatGPT, Gemini, Google AI) and understandably, they indicated the reason why they use them: primarily for vocabulary and assignment support (See Q9).

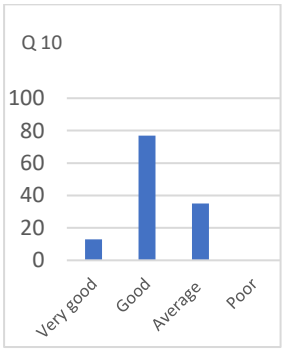
Fig.8 - Time refering to AI tools for SLA



42% said that they referred to AI tools daily. 49% said that they referred to AI tools at weekends and in their free time. Only 02% confirmed that they did not refer to AI tools for the subject at all.

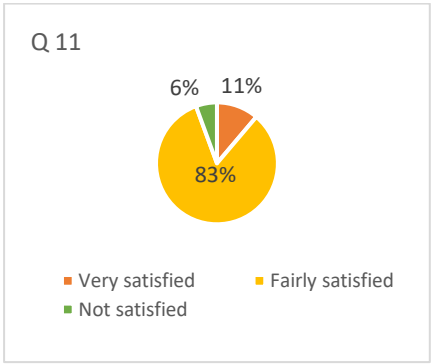
4.4. Satisfaction and Impact (Q10, Q11, Q12)

Fig.9 - Rating the responses by AI Tools



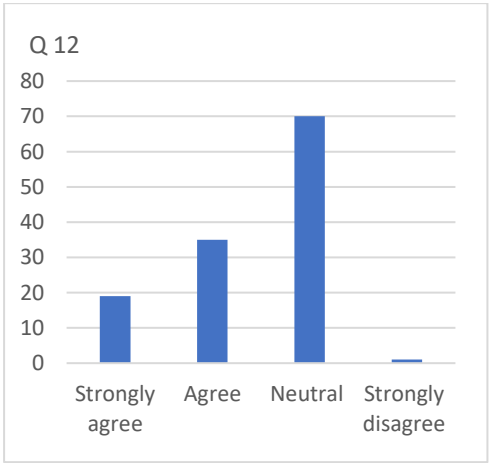
72% claimed that AI responses were good (with 10% - very good. 62% stated that AI feedback was good enough). 28% claimed it was average. No respondents claimed that AI feedback was poor.

Fig.10 - Satisfaction referring to AI Tools



94% felt happy with AI references (with 11% - extremely satisfied and 83% - fairly satisfied). Only 6% of the respondents claimed “not satisfied”.

Fig.11 - Levels of improvement when using AI Tools/ Apps



43% of them said that they have impvoved their learning when referring to AI tools (with 15% "strongly aggre" and 28% "agree"). 70 students accounting for 56 % stayed "neutral".

4.5. Findings of Q14 (Qualitative measure)

Themes	Initial Codes	Key Quotes by respondents as learners
Pedagogy & Instruction	Practice, tips, theoretical balance	"Theory needs to be applied to practice", "provide some tips for better understanding".
Curriculum Design	Challenge, compatibility	"SLA course material is truly challenging for first-year students".
Material Quality	Academic terms, lenghthy textbooks	"Textbook is too long and wordy", "many academic terms that are difficult to grasp".
AI Integration	AI overuse, critical thinking	"Reduce the overuse of AI to develop students' thinking skills".

5. Discussion

5.1. Quantitative Analysis (Q1 - Q13)

5.1.1. The "Utility vs. Value" Paradox

The "Utility vs. Value" paradox reveals a significant gap: while 93% of students acknowledge AI's competitive advantage, only 84% view Second Language Acquisition (SLA) as a necessary subject. This discrepancy suggests that many learners treat AI as a "shortcut" to bypass the cognitive struggle required to master complex theories. Instead of employing AI as a cognitive tool to deepen their understanding, students may rely on it as a compensatory tool simply to finish tasks faster. This trend risks prioritizing efficiency over genuine intellectual growth, potentially undermining the long-term internalisation of core SLA concepts.

5.1.2. The Performance-Learning Gap

The study reveals a critical "Performance-Learning Gap" - Students use it and they like the quality (Q10), but they aren't sure if it is making them "better" students yet. A standout contradiction in the data lies in the discrepancy between student satisfaction and perceived learning growth (Q11 vs. Q12). While a vast majority of respondents expressed satisfaction with the quality of AI-generated responses, over half (56%) remained "Neutral" regarding whether these tools significantly improved their actual mastery of Second Language Acquisition (SLA) core knowledge. This suggests a phenomenon that could be termed as "Illusion of Competence."

To be more specific, students may experience high user satisfaction because AI reduces the cognitive load required to complete immediate tasks, such as defining terminology or structuring assignments. However, this ease of task completion does not necessarily translate into deep cognitive processing. In the context of SLA, where internalizing complex theories require active mental struggle, the assistance provided by AI may be bypassing the very "desirable difficulties" essential for long-term retention. Consequently, without intentional pedagogical scaffolding to ensure AI tools promote active engagement rather than passive reliance, learners may achieve the desired end-product without developing the internal linguistic mental models required for long-term proficiency.

5.1.3. The Role of the Instructor in the AI Era

Recent findings highlight that 94% of students still view teachers' lectures as their primary motivator, significantly surpassing the influence of AI and smartphones. This underscores that even for "digitally skilled learners", the human element remains the core driver of academic engagement. Consequently, AI must be viewed strictly as a supplement rather than a replacement for professional guidance. Instructors play a vital role in interpreting complex SLA concepts, providing nuanced pedagogical insights and personal connections that technology cannot replicate effectively.

This survey aligns with Kostka & Toncelli 's points. It is undeniable that while ChatGPT offers many advantages, it is not a replacement for a qualified human teacher. Effective integration of ChatGPT into an English language classroom would require thoughtful planning and a blended approach, combining technology with human interaction, guidance, and feedback. A teacher's expertise in understanding individual student needs, offering encouragement, and fostering a supportive learning community is irreplaceable and would complement the benefits offered by ChatGPT (Kostka & Toncelli, 2023).

Although technologies may be powerful and advanced, they cannot function without human cooperation (Dziak, 2024). It has clearly been shown that teachers can promote and sustain higher motivation levels by helping learners to identify short-term goals and to reflect on their progress and achievements (Cook & Singleton, 2014).

5.1.4. Implications for Pedagogy and Curriculum Design

The findings of this study necessitate a shift in how SLA is taught within the English Department at HUFLIT. Since students demonstrate high digital literacy but uncertainty regarding AI's impact on their learning growth, the curriculum should move away from traditional "output-based" assignments that AI can easily replicate. Instead, the following implications should be considered. First, it is Integration of AI-Reflective Tasks: Rather than banning AI, lecturers should design tasks where students must use AI to generate an SLA analysis and then critically evaluate that analysis using core textbook theories. This forces the "Neutral" group to engage in the deep processing they feel are currently missing. Second, it is Assessment Reform: With 93% of students recognizing the "AI advantage", assessment should pivot toward in-class, process-oriented evaluations - such as the debate skills already emphasized in the curriculum - where the human element of spontaneous linguistic interaction remains central.

5.2. Qualitative Analysis (Q14)

Qualitative analysis of the open-ended responses reveals four primary themes concerning the enhancement of teaching and learning quality.

First, regarding pedagogical approaches, students emphasized the necessity of bridging the gap between theory and practice. Respondents noted that "theory needs to be applied to practice" and suggested an increase in practical contact hours.

Second, the curriculum - specifically the SLA course - was perceived as overly broad and theoretical, posing significant challenges for first-year students. Consequently, learners expressed a need for instructional tips and clearer guidance from instructors to facilitate better comprehension.

Third, the quality of learning materials was identified as a major pedagogical barrier. A majority of respondents characterized the current textbook as "too long and wordy", noting that it contains numerous "academic terms that are difficult to grasp". This suggests an urgent need to streamline content and adopt materials that are more accessible and tailored to the proficiency levels of novices and appropriateness.

Finally, a concern regarding technology integration emerged; participants recommended mitigating the overuse of AI to prioritize the development of students' independent thinking skills. These findings collectively underscore the requirement for a strategic realignment of both course materials and instructional methods to better suit the developmental needs of first-year students.

6. Conclusion

6.1. Evaluation

The survey clearly reflects students' perceptions and positive attitudes towards the subject, toward the integration of AI in SLA, predominant AI tools and usage patterns exhibited by these students in their language learning routines. In other words, it reveals the trend of students' strong engagement with digital technology, their openness but alertness when using digital tools, and full awareness of the necessary factors to improve the subject, especially the textbook as learning materials. These results are an important basis for improving the SLA curriculum and effectively combining traditional methods and modern technology in training.

6.2. Recommendations

6.2.1. Recommendations from survey respondents

Researchers have focused on how learners react to how they engage with the target language, including how they respond to the teacher's approach, to the learning materials to which they have access (Cook & Singleton, 2014). The feedback on question 14 showed the students' recommendations on the learning materials they had just learned, including three main points:

(1) Practicality: Theory needs to be applied to practice; more practical hours are needed.

(2) Content: What they had learned seemed too theoretical and difficult to grasp. The content should be streamlined to suit the learning needs of first-year students. The instructor needs to provide some tips for better understanding.

(3) SLA course material: It is truly challenging for first-year students; The textbook is too long and wordy. The textbook needs improvement; The material should be less dry, as it contains many academic terms that are difficult to grasp.

6.2.2. Recommendation from author

In the era of high-tech and AI, teachers must proactively develop AI literacy to avoid becoming outdated. Integrating these tools allows educators to personalize learning, automate

administrative tasks, and create engaging content. Proficiency in AI doesn't just simplify workloads; it empowers teachers to provide more effective, data-driven support, ensuring they remain relevant in a modern classroom. AI tools may help create interactive input in SLA.

6.3. Limitations

Although the study has achieved its goals, there are still some shortcomings. The study presents the inherent limitations of focusing on a specific learner group. In other words, the investigation was a minor study, and its findings were applied to only 2 classes of The Faculty of Foreign Languages of Ho Chi Minh University of Foreign Languages - Information Technology. Thus, it may generally not represent universities all over the country.

The survey was completed in a short amount of time. The students had roughly 15 minutes to complete the questions, which may influence their reflections. Due to the limited availability of time, only some factors were carefully analyzed, other facets maybe not be properly speculated.

6.4. Implications

These findings not only open up directions for further research but also provide a practical basis for educators to improve their curricula, effectively combining traditional methods and modern technology, in order to optimize the training process.

The students' responses reveal several significant implications regarding their attitudes, learning strategies, and expectations concerning the integration of AI in education, particularly for the SLA subject.

(1) Desire for Self-Reliance: Students value their own intellectual effort and want to avoid becoming passive recipients of AI-generated content, suggesting a motivation to develop their own skills rather than solely relying on external tools.

(2) Recognition of AI's Limitations and Strengths: Students understand that AI is a "suggestion, not a replacement".

(3) Proactive and Adaptable Mindset Towards Technology: Instead of resisting AI, students are open to embracing it. They see AI as an inevitable part of the future workforce and want to be equipped with the skills to use it effectively, aligning with development trends.

(4) Curriculum Redesign: Move away from "wordy" textbooks toward more interactive, AI-assisted practical workshops.

(5) Critical AI Literacy: A framework for "Guided AI Use" should be set up where the instructor sets boundaries on when to use ChatGPT versus when to use human critical thinking. Students need to learn how to appropriately and ethically incorporate AI-generated content into their own work (Tseng & Warschauer, 2023).

AI cannot ever replace us in these essential aspects of our work as educators, yet it might make our work more interesting, engaging, and innovative (Kostka & Toncelli, 2023).

References

Aktay, S., Gök, S., & Uzunoğlu, D. (2023, July 31). ChatGPT in education. *TAY Journal, 7*(14), 1–18. https://dergipark.org.tr/en/pub/tayjournal/issue/76658/1259832

AIT. (2023, January 17). *6 ứng dụng của AI – Trí tuệ nhân tạo trong giáo dục*. AIT School

of Management. https://som.edu.vn/6-ung-dung-cua-tri-tue-nhan-tao-trong-giao-duc/

An, D. T. (2023, April 3). *ChatGPT: Lợi ích và những cảnh báo trong giáo dục*. Vietnam Innovation of General Education Foundation. Tạp chí Điện tử AUTOMATION. https://tudonghoangnay.vn/chatgpt-loi-ich-va-nhung-canh-bao-trong-giao-duc-240.html

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI, 7,* 52-62. https://doi.org/10.61969/jai.1337500

Benati, A. G., & Angelovska, T. (2016). *Second language acquisition: A theoretical introduction to real-world applications*. Bloomsbury Academic. ISBN 9780567079756.

Chun, D. (2016). *The role of Technology in SLA Research* - Language Learning & Technology, June 2016, Volume 20, Number 2, pp. 98–115 https://www.researchgate.net/publication/304045358_The_role_of_technology_in_SLA_research

Cook, V., & Singleton, D. (2014). *Key topics in second language acquisition*. CPI Group (UK Ltd).

Ccelingo. (2023). *The role of cultural nuances in language translation*.

https://www.accelingo.com/cultural-nuances-in-language-translation-and-their-role-in-professional-language-services/#h-contextual-adaptation

Davis, F. D., Jr. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology). DSpace@MIT. http://hdl.handle.net/1721.1/15192

Dziak, M. (2024). *Technology acceptance model (TAM)*. EBSCO. https://www.ebsco.com/research-starters/technology/technology-acceptance-model-tam

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course* (5th ed.). Routledge.

Godwin-Jones, R., O'Neill, E., Ranalli, J. (2024), Chapter 2 - Integrating AI Tools into Instructed Second Language Acquisition, Technology for Second Language Learning. https://www.iastatedigitalpress.com>pdf

Grand View Research. (2023). *AI in education market size & share – Industry report 2030*. https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report

Henshaw, F. G., & Hawkins, M. D. (2022). *Common ground: Second language acquisition theory goes to the classroom*. Hackett Publishing Company, Inc.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237-257. https://doi.org/10.1111/jcal.12610

Jones, L. C. (2014). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 621–646). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.031

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3*(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115

Karanjakwut, C., & Charunsri, K. (2025). Transforming AI chatbots for brainstorming teaching techniques of process writing. *Malaysian Online Journal of Educational Technology, 13*(1), 1–18. https://www.researchgate.net/publication/387793430_Transforming_AI_Chatbots_for_a_Brainstorming_Teaching_Technique_of_Process_Writing

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *The Electronic Journal for English as a Second Language*, 27(3). https://doi.org/10.55593/ej.27107int

Mutiara, A., Maulida, D., & Indriani, L. (2024, December 29). Challenges of perplexity as artificial intelligence to brainstorm ideas in academic writing. *Journal of English Language Teaching, Literature, and Linguistics, 12*(4), 1845–1856. https://itell.or.id/journal/index.php/itelljournal/article/view/24

Najimova, P. (2024). 'Translating emotion: How cultural nuances affect the interpretation of emotional expressions in literature'. SAECULII YK. https://conference.uzswlu.uz/conf/article/view/250?articlesBySimilarityPage=51

Oanh, P. (2023, March 31). *Ứng dụng ChatGPT vào giáo dục: Phản ứng của các trường Đại học*. [Source Link Missing in Original Image].

Rahman, M. M., Watanabe, Y., & Watanabe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), Article 5783. https://doi.org/10.3390/app13095783

Rasul, T., Nair, S., Kalendra, D., Robin, M., De Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching, 6*(1). https://doi.org/10.37074/jalt.2023.6.1.29

Reinders, H. (2018). *Technology and autonomy*. John Wiley & Sons, Inc. https://www.researchgate.net/publication/373241924_Technology_and_Autonomy

Saville-Troike, M., & Barto, K. (2017). *Introducing second language acquisition* (3rd ed.). Cambridge University Press.

TARI AI. (2024, May 9). *AI tools by TARI*. HUFLIT LMS Moodle. https://tari.huflit.edu.vn/course/view.php?id=95

The Vietnam Foundation. (2025, March 28). *AI trong giáo dục: Hiểu AI để dùng AI hiệu quả*. VNF News. https://vnfoundation.org/ai-trong-giao-duc-hieu-ai-de-dung-ai-hieu-qua/

Tseng, W. T., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning, 3*(2), 258-262. https://doi.org/10.1515/jccall-2023-0008

VanPatten, B., & Benati, A. G. (2020). *Key terms in second language acquisition* (2nd ed.). Bloomsbury Publishing Plc.

VanPatten, B., Smith, M., & Benati, A. G. (2020). *Key questions in second language acquisition* (1st ed.). Cambridge University Press.

Warschauer, M. (1996). *Computer assisted language learning: An introduction*. In S. Fotos (Ed.), *Multimedia language teaching*. Tokyo.

Zebua, J. A. Z., & Katemba, C. V. (2024). Students' perceptions of using the OpenAI ChatGPT application in improving writing skills. *Journal of Language and Literature Studies*, *4*(1), 110–123. https://doi.org/10.36312/jolls.v4i1.1805

Bio:

She earned her M.A in Applied Linguistics from La Trobe University Australia. She has over 30 years of English teaching experience at university level. Her first job was working at The Office of International Affairs, HCM University of Technology as an interpreter and translator. Her second job was teaching ESP at HCM University of Technology. Then she has so far worked at Huflit as a lecturer in Faculty of Foreign Languages. Her interest includes teaching English Linguistics and ESP. As an educator, she welcomes AI integration, but she absolutely appreciates the combination between traditional teaching approach and new ones.

Appendix

SURVEY QUESTIONNAIRE

Question 1: Do you think the subject of Second Language Acquisition (SLA) is necessary for your curriculum?

a) Very necessary

b) Necessary

c) Not necessary

d) Completely unnecessary

Question 2: How would you rate your understanding of the core knowledge of SLA?

a) Fairly good

b) Average

c) Below average

d) Not at all

Question 3: In your opinion, which of the following factors help students be more motivated to study SLA? *(You may choose multiple options)*

a) Internet-connected devices (computer, laptop, iPad, etc.)

b) Smartphones

c) Application software (Apps)

d) Textbooks and reference books

e) Various types of dictionaries

f) Teacher's lectures in class

g) Other:

_____

Question 4: When you need to understand the core knowledge of SLA, do you often consult AI/AI Apps? (e.g., ChatGPT, Gemini, Google AI, Online Dictionary, etc.)

a) Regularly

b) Occasionally

c) Rarely

d) Never

Question 5: Since when have you known about AI tools and its applications in education?

a) Less than 1 year

b) 1 to 2 years

c) More than 2 years

d) More than 3 years

Question 6: Could you please list the AI tools/apps that you use for the SLA subject?

_____

_

Question 7: How would you rate your proficiency in using AI tools/apps?

a) Very good

b) Good

c) Not very good

d) Not at all

Question 8: How much time do you spend using AI tools for this subject?

a) Daily

b) Weekends or free time

c) Only when needed

d) No time spent on AI tools

Question 9: For what purposes do you use AI and Apps when studying SLA?*(You may choose multiple options)*

a) Checking vocabulary and terminologeies of the subject

b) Generating ideas for class assignments

c) Supporting assignments on Moodle

d) Checking answers to exercises

e) Focusing on core lessons

f) Other purposes: _____

Question 10: How would you rate the responses provided by these AI tools/apps?

a) Very good

b) Good

c) Average

d) Poor

Question 11: Are you satisfied with the suggestions provided by AI tools/apps for this subject?

a) Very satisfied

b) Fairly satisfie

c) Not satisfied

Question 12: In your opinion, your learning has improved significantly when using AI tools/**apps?**

a) Strongly agree

b) Agree

c) Neutral

d) Disagree

e) strongly disagree

Question 13: Which group of people do you think has an advantage in learning?

a) People who use AI tools/apps effectively

b) People who do not use AI tools/apps

c) Other opinions:

_____

Question 14: Please provide any suggestions you have for the SLA course.

_____

_

Thank you very much for participating in this survey.

#49 Developing Intercultural Pragmatic Competence for Real Life Interaction through Role-Play: Teaching the Speech Act of Requests in Vietnamese English Communication Classes in the Transition from EFL-Oriented to ESL-Oriented Pedagogy

Nguyen Hoang Vuong Anh, Nguyen Tan Khoa

Abstract

In recent years, English education in Vietnam has increasingly emphasized communicative competence and the use of English for interaction as the country shifts toward a more ESL-oriented context. However, classroom instruction still largely prioritizes grammatical knowledge and linguistic accuracy over pragmatic aspects of communication. Within communicative competence, the ability to perform speech acts appropriately is essential, as effective interaction requires sensitivity to social context, power relations, and cultural norms. Among speech acts, requests are particularly important due to their variability across languages and their role in managing interpersonal relationships. Despite this significance, instructional approaches to teaching requests remain underexplored in Vietnam.

This study examines how role-play activities can be used to teach the speech act of requests in Vietnamese English communication classes and how they support the development of students' intercultural pragmatic competence. Adopting a classroom-based case study design, the research involved 41 freshmen enrolled in an English communication course at the Faculty of Foreign Languages, HUFLIT University. Data were collected through classroom observations, recordings of role-play interactions, and students' task performances, and were analyzed using pragmatic discourse analysis.

The findings show that role-play enhances students' ability to perform the speech act of requests appropriately across diverse social contexts and provides empirical evidence for integrating pragmatic, interaction-based instruction. These results highlight the significance of role-play as a pedagogically effective approach in facilitating the transition from grammar-focused EFL practices to a more communicative, socially grounded ESL-oriented approach in Vietnam.

***Keywords*:** intercultural pragmatic competence, role-play, speech acts (requests), ESL-oriented pedagogy

1. Introduction

Effective English communication involves both grammatical accuracy and contextually appropriate language use. This ability, known as pragmatic competence, is especially important when learners move from EFL to ESL settings, where communication occurs more frequently in real-life situations.

In university contexts, students often need to perform tasks such as emailing lecturers, asking for clarification, or requesting assistance. These interactions require an understanding of politeness and social relationships. Nevertheless, classroom practices in many Vietnamese EFL settings continue to prioritize linguistic forms and vocabulary development, leaving relatively little room for the systematic development of pragmatic abilities.

Previous studies have examined learners' request strategies and highlighted the effectiveness of role-play activities in developing pragmatic competence and intercultural awareness. However, little research has explored how interaction-based role-play activities help students improve the pragmatic appropriateness of requests, develop request strategies and mitigating devices, and facilitate the shift from grammar-focused EFL instruction to communicative ESL-oriented learning.

Therefore, further research is needed to explore how role-play activities contribute to developing intercultural pragmatic competence among Vietnamese university students.

To address this issue, the study examines how interaction-based role-play activities facilitate students' use of the speech act of requests and the development of intercultural pragmatic competence. Specifically, it examines the pedagogical implementation and analyzes how these activities facilitate students' transition toward an ESL-oriented communicative approach. To guide this investigation, the following research questions are proposed:

(1) To what extent does the integration of role-play activities improve students' pragmatic appropriateness in performing the speech act of requests?

(2) How do students develop request strategies and use of mitigating devices during role-play activities?

(3) How do students perceive the role of role-play in supporting the shift toward more ESL-oriented communicative practices?

The study first outlines key concepts related to interlanguage pragmatics and intercultural competence. The next section presents the research design, participants, and data collection process. The findings are subsequently presented and analyzed, followed by a discussion of pedagogical implications. The final section summarizes the findings and offers recommendations for future research.

2. Literature Review

2.1 Pragmatic Competence in Interlanguage and Intercultural Pragmatics

2.1.1. Pragmatic Competence

Pragmatic competence is a key aspect of communicative competence. According to Hymes, D. (1972), successful communication depends on both linguistic knowledge and appropriate language use. Based on this idea, Canale, M., & Swain, M. (1980) regarded sociolinguistic competence as an essential element of communicative competence. Later, pragmatic competence was further described as a fundamental component of language proficiency that combines knowledge of language forms with an understanding of social and cultural contexts (Bachman, L. F.,1990). Collectively, these perspectives indicate that effective communication requires a combination of linguistic resources and context-sensitive language use.

More specifically, researchers commonly describe pragmatic competence as the capacity to select and employ language forms that are appropriate for a given communicative situation. According to Leech, G. N. (1983) and Thomas, J. (1983), successful language use depends on selecting language that fits the communicative context. Pragmatic competence combines language knowledge with an understanding of social and cultural expectations. (Kasper, G., & Rose, K. R., 2002). As a result, communication problems may arise when learners focus on grammar without considering social and contextual appropriateness. (Blum-Kulka, S. et al., 1989).

Pragmatic competence is also closely related to culture. The same expression can have different meanings in different cultural and social contexts, which may lead to misunderstanding (Ishihara, N., & Cohen, A. D., 2010). Thomas, J. (1983) argued that communication problems often result from pragmatic failure rather than grammatical errors. For this reason, teaching pragmatics should go beyond language forms and help learners develop intercultural awareness. This can enable them to communicate more effectively and appropriately in different situations.

2.1.2 Interlanguage Pragmatics

Interlanguage pragmatics investigates learners' acquisition and use of pragmatic knowledge. Kasper, G., & Schmidt, R. (1996) define the field as the study of learners' acquisition and use of language in socially appropriate ways. According to Kasper, G., & Rose, K. R. (2002), interlanguage pragmatics explores how learners develop pragmatic competence and apply it in various communicative situations. Similarly, Bardovi-Harlig, K. (2013) argues that learners are required to adapt their language to various communicative contexts..

A substantial body of interlanguage pragmatics research has investigated speech acts including requests, apologies, and refusals, as well as the teaching and assessment of pragmatic competence (Rose, K. R., 2005 & Taguchi, N., 2015). Despite this growing body of research, interactional competence has received less attention, even though it plays a central role in successful communication (Taguchi, N., & Roever, C., 2017). Learners who lack sociopragmatic awareness may have difficulty interpreting social norms and cultural expectations, which can lead to misunderstandings in intercultural interactions. Such difficulties indicate that learners may benefit from explicit pragmatics instruction.

2.1.3 Intercultural Pragmatics

Intercultural pragmatics examines how cultural knowledge and experience shape cross-cultural communication. Kecskés, I. (2014, 2016) explains that people from different cultures may interpret and use language in different ways during communication. McConachy, T. (2022) further argues that learners' understanding and production of meaning are shaped by their previous experiences and cultural knowledge. As a result, speakers often need to negotiate meaning and build common ground to achieve successful communication.

Pragmatic meanings are also influenced by cultural views of politeness, power, and social relationships rather than being universally shared (Spencer-Oatey, H., 2008; Kádár, D. Z., & Haugh, M., 2013). For this reason, intercultural communication requires flexibility and the ability to mediate between different perspectives. Liddicoat, A. J. (2022) emphasizes the importance of intercultural mediation, while Byram, M. (2011) and McConachy,T. (2018) describe language learners as intercultural speakers who adapt their language use to different contexts instead of simply copying native-speaker norms. These perspectives highlight the need to develop both pragmatic competence and intercultural awareness in language education.

2.2 Requests as Speech Acts in Pragmatics

Requests are important speech acts because they may pressure the listener and threaten their social image (Ellis, R., 2003). Therefore, producing requests requires careful attention to contextual variables such as power relations, social distance, and degree of imposition. In addition, speakers must select appropriate strategies that vary in terms of directness and politeness.

Furthermore, research on request perception shows that indirect requests are highly context-dependent. In particular, Taguchi, N. (2015) demonstrate that hearers rely heavily on

contextual cues to interpret intended meaning. Moreover, advanced learners may develop native-like perceptions of politeness, although their judgments are still influenced by instruction and learning experience (Taguchi, N., & Roever, C., 2017).

In addition, research suggests that request strategies vary across cultures because of different sociocultural expectations (Ogiermann, E., 2009 & Dong, X., 2009). As a result, learners with greater exposure to the target language and authentic interaction are generally more capable of producing contextually appropriate requests (Han, S., 2005).

2.3 Learners' Pragmatic Performance in Using Requests

Braj B. Kachru's three-circle model classifies countries according to the role of English as an inner-circle, outer-circle, or expanding-circle language (Kachru, B. B., 1982 & Crystal, D., 2003). In this regard, the distinction is important because learning environments strongly influence pragmatic development.

Specifically, learners in ESL contexts are exposed to authentic communication and therefore tend to develop more contextually appropriate request strategies (Bardovi-Harlig, K. & Dörnyei, Z., 1998 & Schauer, G. A., 2006). In contrast, EFL learners often have limited opportunities for real-life interaction and depend more heavily on classroom input. As a result, they are more likely to experience stronger L1 influence and less appropriate request performance (Ishihara, N., & Cohen, A. D., 2010).

Furthermore, Austin's Speech Act Theory provides an important framework for studying requests. According to Austin, J. L. (1962), language is used not only to convey information but also to perform actions such as requesting or apologizing. Building on this idea, Searle, J. R. (1979) further distinguishes between direct and indirect speech acts, emphasizing that indirect requests are often preferred because they reduce imposition and maintain politeness. Similarly, Brown, P., & Levinson, S. C. (1987) highlight the importance of indirectness in English-speaking contexts.

In addition, cross-cultural studies compare request strategies across languages and cultures, thereby revealing differences in directness and politeness (Blum-Kulka, S., et al., 1989). By contrast, interlanguage studies focus on how learners acquire and use request strategies in a second language while considering factors such as proficiency, learning environment, and L1 influence (Kasper, G., & Dahl, M., 1991).

2.4 Pragmatics Instruction: Theoretical and Empirical Perspectives

Research on pragmatic instruction is generally informed by both cognitive and sociocultural perspectives. From the cognitive perspective, DeKeyser, R. M. (2007) emphasize the roles of noticing, attention, and explicit instruction in developing pragmatic awareness. In contrast, from the sociocultural perspective, Vygotsky, L. S. (1978) stress the importance of interaction and collaborative meaning-making in language learning.

More recently, these perspectives have been viewed as complementary rather than contradictory. In particular, explicit instruction helps raise learners' metapragmatic awareness, whereas interaction promotes practice and the automatization of pragmatic knowledge (Ishihara, N., 2010 & Taguchi, N., 2015).

Furthermore, empirical studies support the effectiveness of explicit instructional techniques such as metapragmatic explanation and role-play in improving learners' use of request strategies (Takahashi, S., 2010). At the same time, authentic input and interaction play an important role in developing sociocultural understanding and communicative fluency

(Taguchi, N., 2015). However, limited exposure to authentic communication in EFL settings may hinder pragmatic development (Kasper, G., & Rose, K. R., 2002). Therefore, combining explicit instruction with interaction-based activities is widely considered an effective approach to teaching pragmatics.

3. Methodology

3.1 Research Setting

This research was carried out at the Faculty of Foreign Languages, HUFLIT, a private university in southern Vietnam.

The research took place in an English communication course for first-year students, which is a compulsory subject aimed at enhancing learners' speaking and interactional skills in real-life contexts. The course focuses on developing students' ability to use English effectively in everyday communication, incorporating interactive teaching methods such as role-play, group discussions, and situational simulations. In particular, role-play activities are regularly employed to simulate authentic communicative situations, allowing students to practice language functions such as making requests, giving suggestions, and expressing opinions.

The course is typically organized into weekly sessions that combine both theoretical input and practical activities. These activities are designed not only to improve linguistic competence but also to raise students' awareness of sociocultural norms and pragmatic conventions in English communication.

3.2 Participants

The study involved 41 first-year students in an English communication course. Convenience sampling was used, with participants selected based on availability at HUFLIT. Although the study was conducted in a single institutional context, the classroom reflected common characteristics of Vietnamese EFL communication classes that are increasingly incorporating ESL-oriented communicative practices.

3.3 Research Instruments

To examine students' use of request strategies and the development of intercultural pragmatic competence, this study adopted a mixed-methods approach, employing multiple instruments: a Written Discourse Completion Task (WDCT), a questionnaire, role-play tasks, and classroom observations. Employing multiple sources of evidence strengthened the trustworthiness and credibility of the research outcomes.

The WDCT was used to elicit written request strategies across six scenarios reflecting variations in power, distance, and imposition (Brown, P., & Levinson, S. C., (1987). A questionnaire (25 items) investigated students' perceptions of role-play and pragmatic development, combining Likert-scale items (e.g., pedagogical effectiveness, pragmatic competence, intercultural awareness, psychological factors) and multiple-choice questions on communicative behavior and L1 influence.

Role-play tasks served as the primary source of spoken data, requiring students to perform requests in interactive contexts based on similar sociopragmatic variables. Their interactions were audio-recorded for analysis. In addition, classroom observations were conducted to capture students' engagement and interaction patterns, supported by field notes.

3.4 Procedure

Data were collected during regular English communication classes through multiple complementary stages. First, a WDCT with six scenarios was administered online to elicit written request strategies in controlled contexts, providing baseline data.

Second, role-play activities were implemented to simulate real-life interactions. Students performed in pairs or groups, and their interactions were audio-recorded to capture authentic spoken data. Classroom observations were conducted concurrently, with field notes documenting engagement and interaction patterns.

Finally, a 25-item questionnaire was administered to gather students' perceptions of role-play and its contribution to their intercultural pragmatic development. Responses were collected anonymously to ensure reliability.

Overall, the use of multiple data sources such as written (WDCT), spoken (role-play), observational, and self-reported enabled triangulation and a comparative analysis of students' pragmatic performance across task types.

3.5 Data Analysis

Data from multiple instruments were analyzed using both qualitative and quantitative approaches to examine students' use of request strategies and their intercultural pragmatic competence. First, the WDCT responses were exported to Excel and coded deductively based on Blum-Kulka and Olshtain's framework, focusing on levels of directness and politeness features; subsequently, the data were quantified to identify patterns across contexts. Next, role-play recordings were transcribed and analyzed using pragmatic discourse analysis to examine strategy use, directness, and contextual appropriateness in real-time interaction. Meanwhile, questionnaire data were analyzed quantitatively, with Likert-scale items calculated in terms of frequencies, percentages, and mean scores, and multiple-choice items examined descriptively to identify trends in learners' perceptions and strategy use. In addition, classroom observation notes were analyzed qualitatively to provide supporting evidence and triangulate the findings.

Overall, the integration of multiple data sources enhanced the validity and reliability of the study.

4. Results

4.1. Vietnamese EFL Learners' Realization of Speech Act of Requests: Findings from the WDCT

The quantitative analysis of the WDCT provides a comprehensive baseline of the participants' pragmatic competence. The data were categorized based on the directness levels and specific request strategies following the Cross-Cultural Study of Speech Act Realization Patterns (CCSARP) framework.

4.1.1. Overall Distribution of Request Directness

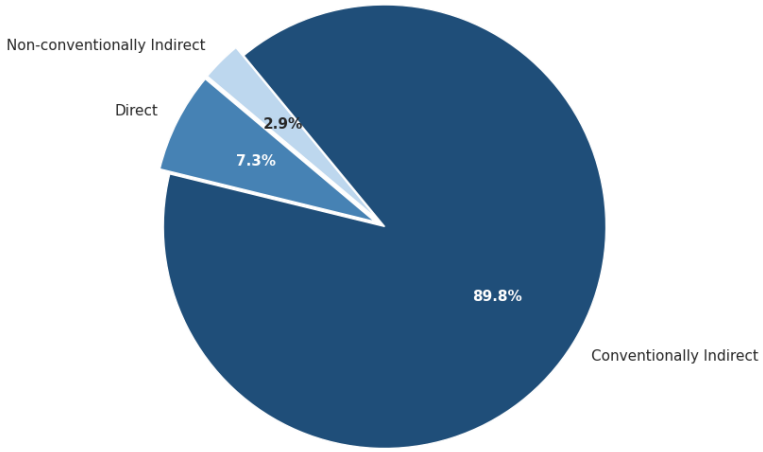


Figure 1. *Distribution of Request Directness in WDCT (N=41)*

As presented in Figure 1, the participants demonstrated a clear preference for Conventionally Indirect strategies, which accounted for the vast majority of the total requests (89.8%). In contrast, Direct strategies were used far less frequently (7.3%), while Non-conventionally Indirect strategies (Hints) were the least frequently employed (2.9%). The results indicate that participants generally recognized the importance of politeness and indirectness when making requests in English. However, the overwhelming reliance on conventionally indirect forms also indicates a tendency toward formulaic language use, with limited variation in pragmatic strategy selection across contexts. While the learners appeared capable of producing pragmatically acceptable requests in many situations, the findings also suggest that their pragmatic repertoire remained relatively restricted.

4.1.2. Situational Variation in Request Strategies

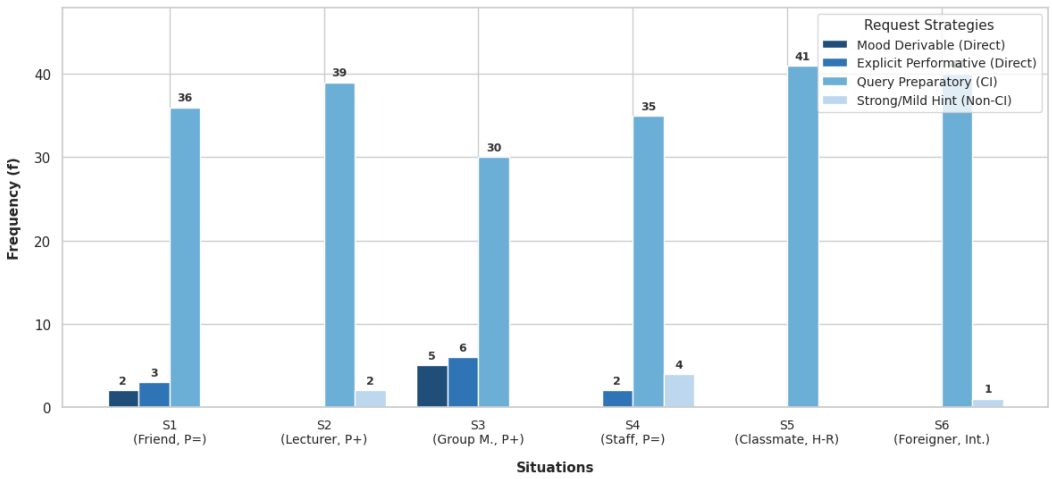


Figure 2. *Distribution of Request Strategies across Six Situations (N=41)*

Figure 2 demonstrates that the Query Preparatory strategy was the most dominant across all six scenarios, maintaining a near-absolute presence in Situations 5 and 6 (100% and 98% respectively). This indicates that when faced with high imposition (S5) or intercultural contexts (S6), learners almost exclusively resort to safe, conventionalized questions about ability or willingness (e.g., "Can you...").

However, a notable deviation in directness levels was observed in Situation 3 (Group Leader to Member, P+), where Direct strategies (Mood Derivable and Explicit Performative) reached their highest frequency at 27%. This shift suggests that learners are sensitive to the Power (P) variable; they tend to adopt a more authoritative and blunt tone when they perceive themselves in a superior social position.

In contrast, Situation 2 (Lecturer, P+) and Situation 4 (Staff, D+) elicited the most consistent use of polite, indirect formulas. Interestingly, while Non-conventionally Indirect strategies (Hints) remained the least used overall, they appeared more frequently in Situation 4 (10%), as learners attempted to use subtle remarks to address a stranger in a service context.

Overall, the data in Table 2 reinforces the finding that while Vietnamese EFL learners at HUFLIT attempt to calibrate their requests according to social variables like Power and Distance, their pragmatic repertoire remains constrained. They rely heavily on a "one-size-fits-all" indirectness for most scenarios, only occasionally venturing into direct commands when the social hierarchy permits.

4.2. Vietnamese EFL Learners' Pragmatic Performance in Real-Time Interaction: Findings from the Role-Play Recordings

4.2.1. Patterns of Pragmatic Performance in Role-Play Interaction

Analysis of four role-play recordings revealed a consistent pattern in students' pragmatic performance during real-time interaction. Overall, the findings indicate that while students demonstrated awareness of polite request strategies in written tasks, they experienced notable difficulty maintaining such politeness when engaged in spontaneous spoken communication.

One prominent trend was a shift from indirect to more direct request strategies. In WDCT, students typically produced longer and more mitigated forms, such as "I wonder if you could help me…". However, in the recordings, these were frequently replaced by shorter and more direct expressions, including "Give me" or "Can I borrow". This shift suggests that under time pressure, students prioritized efficiency and immediate comprehensibility over pragmatic appropriateness.

4.2.2. Key Challenges in Real-Time Request Production

The analysis also revealed evidence of first language (L1) pragmatic transfer. In several instances, students appeared to rely on Vietnamese syntactic and pragmatic patterns when formulating requests in English. For example, the utterance "Give me your notebook" reflects a direct translation of a common Vietnamese request form ("Cho mình mượn"), without the use of mitigating devices typically expected in English. Although not intended to be impolite, such forms may be perceived as abrupt in English-speaking contexts.

In addition, the recordings showed clear signs of processing difficulty. Across all four interactions, students frequently produced hesitation markers such as "uhm" and "ah", along with pauses ranging from two to four seconds. These features suggest a cognitive load

associated with simultaneously managing lexical retrieval, grammatical encoding, and pragmatic decision-making during real-time speech production.

Finally, students demonstrated difficulty handling unexpected responses from their interlocutors. When faced with refusals or conditional replies (e.g., a partner indicating that they also needed the item), many students struggled to adapt their requests. Instead of employing supportive moves such as providing justification or negotiation (e.g., "I only need it for a moment" or "I'll return it immediately"), they often repeated their initial request or responded minimally without further elaboration.

In summary, the findings from the four role-play recordings suggest a gap between students' declarative knowledge of politeness strategies and their ability to apply this knowledge in real-time interaction. Under communicative pressure, their language use became more direct and was strongly influenced by L1 pragmatic norms.

4.3. Learners' Pragmatic Perception and Pedagogical Evaluation: Findings from the Questionnaire

4.3.1 Students' Perceptions toward Role-play Activities

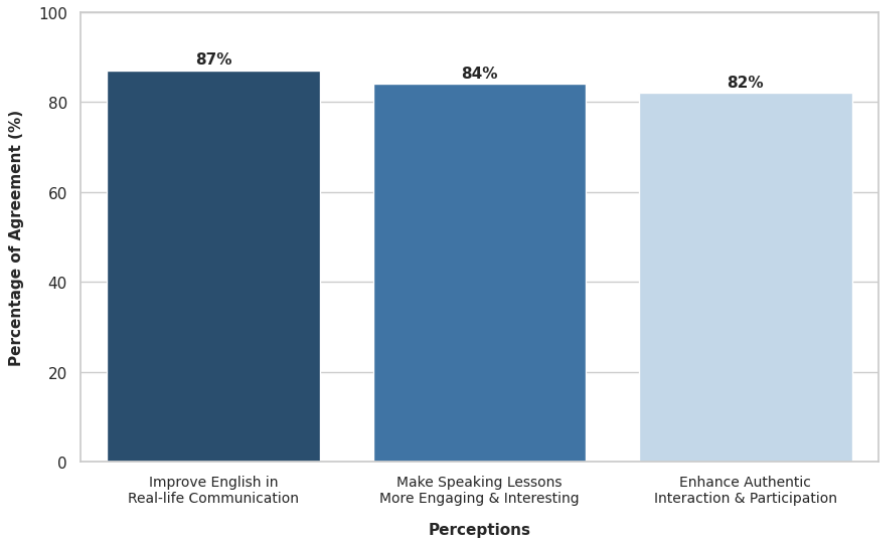


Figure 3. *Students'Perceptions toward Role-play Activities*

The questionnaire results revealed generally positive attitudes toward the use of role-play activities in English communication classes. Most participants agreed that role-play created a more engaging and interactive learning environment.

As illustrated in Figure 3, approximately 87% of the participants agreed that role-play activities improved their ability to use English in real-life communication, indicating that students perceived role-play as beneficial for communicative development. Similarly, around 84% of the respondents reported that role-play activities made speaking lessons more engaging and interesting. In addition, about 82% of the students believed that role-play activities enhanced authentic interaction and active classroom participation. These findings suggest that role-play was viewed as an effective communicative approach that facilitated meaningful language use in classroom contexts.

Overall, the findings demonstrated that students perceived role-play activities positively and considered them beneficial for improving classroom interaction, learner engagement, and English communication competence.

4.3.2 Role-play and Intercultural Pragmatic Development

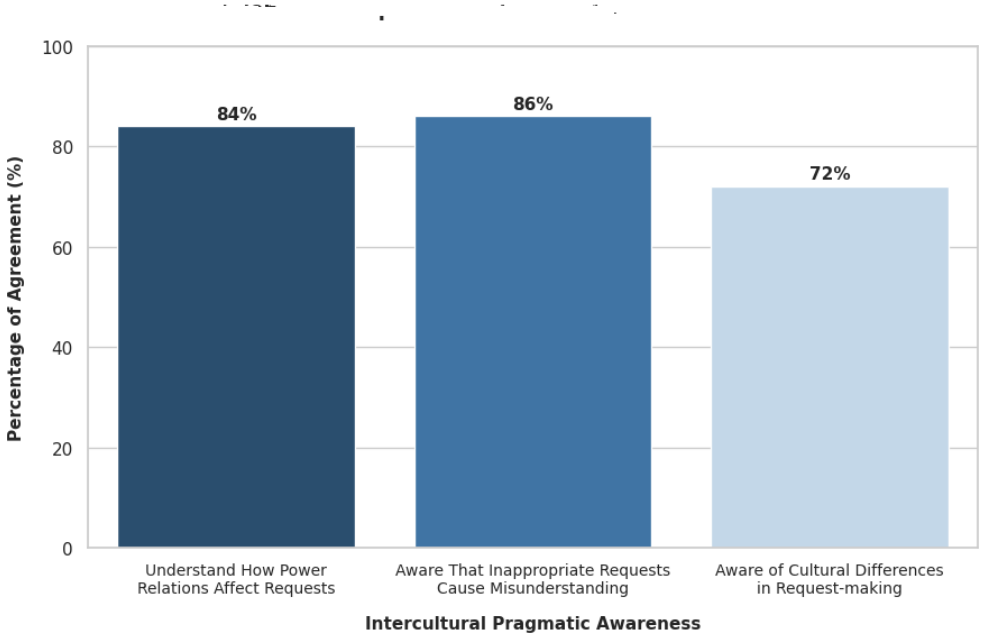


Figure 4. *Students' Perceptions toward Intecultural Pragmatic Development*

The findings revealed that role-play activities contributed positively to students' intercultural pragmatic development. Many participants demonstrated awareness of how social and cultural factors influence communication when making requests in English.

As shown in Figure 4, approximately 84% of the participants agreed that they understood how power relations affect requests, indicating that students recognized the importance of social status and hierarchy in communication. In addition, around 86% of the respondents reported awareness that inappropriate requests may cause misunderstanding in intercultural communication. Students also showed awareness of cultural differences in request-making although the level of agreement for this item was slightly lower, with approximately 72% of participants expressing positive responses. This finding suggests that while students had developed intercultural pragmatic awareness, additional exposure to authentic intercultural communication may still be necessary.

Overall, the findings indicate that role-play activities supported the development of intercultural pragmatic competence by helping students understand contextual appropriateness, politeness, and social relationships in English communication.

4.3.3 Pragmatic Awareness in Using Speech Acts of Request

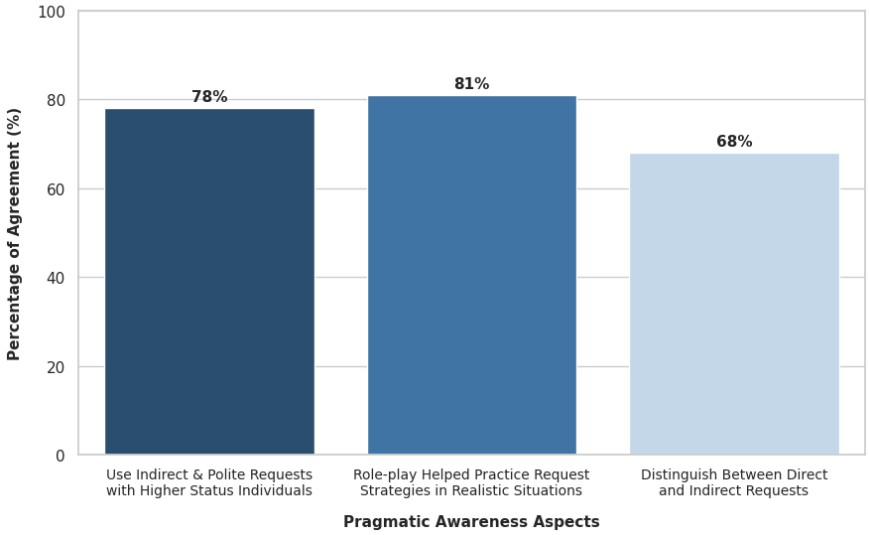


Figure 5. *Students' Pragmatic Awawreness in Using Speech Acts of Request*

The questionnaire results demonstrated that students developed a relatively strong awareness of pragmatic appropriateness when using speech acts of request.

Most participants agreed that requests should be adjusted depending on the context, social relationship, and level of politeness required in communication. As illustrated in Figure 5, approximately 78% of the respondents reported that they tended to use indirect and polite requests when communicating with someone of higher social status.

Additionally, around 81% of the participants indicated that classroom role-play activities helped them practice request strategies in realistic situations. Many respondents also stated that teachers frequently provided feedback regarding the appropriateness and cultural use of requests during classroom interaction. However, the item related to distinguishing between direct and indirect requests received a comparatively lower level of agreement, with approximately 68% of students expressing positive responses. This finding suggests that although students demonstrated general pragmatic awareness, some still experienced difficulty identifying and applying different request strategies accurately.

Overall, the findings suggest that role-play activities positively influenced students' pragmatic awareness and their ability to use speech acts of request more appropriately in English communication.

4.3.4 Emotional and Communicative Effects of Role-play Activities

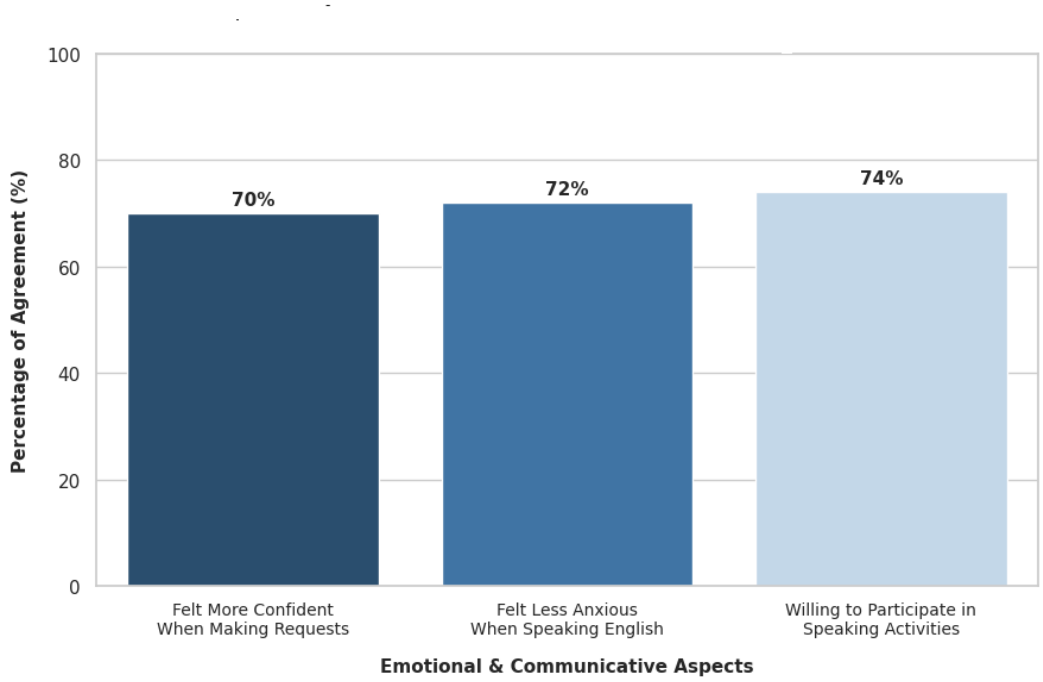


Figure 6. *Emotional and Communicative Effects of Role-play Activities*

The questionnaire findings also revealed several emotional and communicative effects associated with role-play activities.

Many students reported that role-play increased classroom interaction and encouraged participation in speaking activities. The communicative nature of role-play appeared to help students become more actively involved in the learning process.

As shown in Figure 6, approximately 70% of the participants felt more confident when making requests in English. Similarly, around 72% of the respondents reported feeling less anxious when speaking English during role-play activities. In addition, nearly 74% of the students expressed willingness to participate in speaking activities through role-play tasks. These findings suggest that although role-play created a supportive and interactive learning environment, some learners still experienced hesitation and anxiety when communicating in English.

Overall, the findings indicate that role-play activities had positive communicative effects, particularly in promoting classroom interaction and learner engagement. Nevertheless, emotional barriers such as anxiety and lack of confidence remained challenges for certain students and may require additional pedagogical support and speaking practice.

5. Discussion

5.1 Research question 1: To what extent does the integration of role-play activities improve students' pragmatic appropriateness in performing the speech act of requests?

The findings indicate that the integration of role-play activities contributes positively to students' pragmatic appropriateness in performing the speech act of requests. Questionnaire findings presented in Figure 5 revealed that students developed greater awareness of contextual appropriateness when making requests in English. Specifically, approximately 78% of the respondents reported that they tended to use indirect and polite requests when communicating

with someone of higher social status, while around 81% stated that classroom role-play activities helped them practice request strategies in realistic situations.

This perceived improvement is further supported by the WDCT findings illustrated in Figure 1, where Conventionally Indirect strategies accounted for 89.8% of the total requests. The dominance of indirect request forms suggests that learners were generally aware of the importance of politeness and social appropriateness in English communication. In addition, Figure 2 demonstrates that learners adjusted their request strategies according to social variables such as power and distance. For example, students consistently employed more polite and indirect forms when interacting with lecturers and strangers, whereas more direct strategies appeared more frequently when learners occupied higher-status roles, such as group leaders.

From a cultural perspective, role-play activities played an important role in helping learners recognize and adapt their communicative behavior across different social contexts. Through simulated interactions, students were able to practice varying levels of directness and modify their language use according to contextual demands. However, findings from the role-play recordings also revealed that although students demonstrated pragmatic awareness in written tasks, they still experienced difficulty maintaining politeness strategies during spontaneous spoken interaction. Under communicative pressure, many learners shifted toward shorter and more direct request forms, reflecting the influence of processing difficulty and first-language pragmatic transfer.

Overall, these findings demonstrate that role-play activities not only improved students' linguistic performance but also enhanced their pragmatic and intercultural communicative awareness. Nevertheless, the findings also suggest that learners' pragmatic competence remained somewhat constrained in real-time interaction, indicating the need for continued practice and greater exposure to authentic communicative situations.

5.2 Research question 2: How do students develop request strategies and use of mitigating devices during role-play activities?

The findings indicate that students developed greater awareness of polite and indirect request strategies during role-play activities; however, their ability to use more flexible and sophisticated mitigating devices remained relatively limited.

The WDCT findings presented in Figure 2 demonstrate that the Query Preparatory strategy was overwhelmingly dominant across all six situations and reached nearly absolute frequency in high-imposition and intercultural contexts, accounting for 100% in Situation 5 and 98% in Situation 6. These results suggest that students preferred conventionally indirect and "safe" request forms in order to avoid sounding impolite or inappropriate. In addition, Figure 1 revealed that Conventionally Indirect strategies accounted for 89.8% of all requests, further indicating learners' strong reliance on indirectness as a politeness strategy.

Questionnaire findings shown in Figure 5 also revealed that approximately 78% of the participants reported using indirect and polite requests when communicating with someone of higher social status, while around 81% agreed that role-play activities helped them practice request strategies in realistic situations. Nevertheless, the findings also suggest that students' pragmatic repertoire remained somewhat constrained. Although learners demonstrated awareness of contextual appropriateness, many still relied heavily on fixed and formulaic expressions such as "Can you…?" rather than employing a wider range of mitigating devices and supportive moves.

Furthermore, the role-play recordings revealed that under communicative pressure, students frequently shifted from longer and more mitigated forms in written tasks to shorter and more direct expressions in spontaneous interaction. The recordings also showed hesitation markers, pauses, and evidence of first-language pragmatic transfer, such as direct translations from Vietnamese request forms. These findings suggest that students experienced difficulty managing lexical retrieval, grammatical encoding, and pragmatic adjustment simultaneously during real-time communication.

Overall, while students showed improvement in using indirect request strategies and demonstrated greater awareness of politeness and social appropriateness, their ability to soften, modify, and diversify their requests remained limited. This limitation appears to result from both linguistic constraints and continued reliance on first-language pragmatic patterns during spontaneous spoken interaction.

5.3 Research question 3: How do students perceive the role of role-play in supporting the shift toward more ESL-oriented communicative practices?

The findings indicate that students generally perceived role-play activities as supporting a shift from form-focused learning toward more communication-oriented and context-sensitive language use. Questionnaire findings presented in Figures 3, 4, and 6 revealed that learners increasingly valued communicative effectiveness, contextual appropriateness, and interactional awareness in English communication.

Specifically, Figure 3 showed that approximately 87% of the participants agreed or strongly agreed that role-play activities improved their ability to use English in real-life communication, while around 84% reported that role-play made speaking lessons more engaging and interactive. In addition, about 82% believed that role-play activities enhanced authentic interaction and active classroom participation. These findings suggest that learners viewed role-play as an effective communicative approach that encouraged meaningful language use rather than merely focusing on grammatical accuracy.

Furthermore, Figure 4 demonstrated that students developed greater awareness of intercultural and contextual factors in communication. Approximately 84% of the participants agreed that they understood how power relations influence requests, while around 86% recognized that inappropriate requests may cause misunderstandings in intercultural communication. However, only about 72% expressed confidence in understanding cultural differences in request-making, suggesting that although learners had developed intercultural pragmatic awareness, some uncertainty and cultural discomfort still remained.

The role-play recordings also support this perceptual shift. During interactive tasks, students were required to negotiate meaning, respond spontaneously, and adjust their language according to social relationships and contextual demands. Unlike traditional grammar-focused instruction, role-play activities encouraged learners to prioritize communicative effectiveness and immediate interaction over grammatical perfection. Nevertheless, the recordings revealed that many students still experienced hesitation, pauses, and anxiety during spontaneous communication, indicating that the transition toward communicative competence had not yet been fully consolidated.

Overall, these findings suggest that role-play activities played an important role in promoting communication-oriented learning and raising students' awareness of contextual and intercultural factors in language use. Although students demonstrated a clear shift toward more communicative practices, emotional barriers and limitations in intercultural confidence remained challenges that require continued exposure to authentic communicative interaction.

6. Conclusion

In conclusion, the results provide valuable insights into the ways Vietnamese EFL learners at HUFLIT employ request strategies across different communicative contexts. The findings show that although students are aware of social factors such as Power and Distance, their language use is still limited. They often rely on similar indirect forms in different situations and only use more direct expressions when the context clearly allows it. The analysis of the four role-play recordings also shows a gap between what students know and what they can actually do in real communication. When speaking under pressure, they tend to become more direct and are influenced by their first language. This suggests that students are in a transition stage: they have moved beyond focusing only on grammar, but they have not yet developed flexible and natural communication skills.

From a pedagogical perspective, greater emphasis should be placed on integrating role-play activities into communication-oriented instruction. Students are given opportunities through these activities to practice speaking in realistic situations, which they need to think and respond quickly. Through regular role-play practice, students can gradually learn how to adjust their language, use softer expressions, and become more confident in real interactions. To make role-play more effective, teachers should provide clear guidance, model appropriate language, and give feedback not only on grammar but also on how polite and appropriate the students' requests are.

This study has several limitations, including a small sample size and data collected from only a few role-play tasks, so the results may not fully reflect real-life communication. Future research should include more participants and longer observation periods. It would also be useful to examine both teacher readiness and the use of AI tools to support the teaching of request strategies through role-play activities in Vietnamese English communication classrooms.

References

Austin, J. L. (1962). *How to do things with words*. Clarendon Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning, 63*(Suppl. 1), 68–86.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*(2), 233–259.

Blum-Kulka, S., House, J., & Kasper, G. (Eds.). (1989). *Cross-cultural pragmatics: Requests and apologies*. Ablex.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Byram, M. (2011). Conceptualizing intercultural (communicative) competence and intercultural citizenship. In J. Jackson (Ed.), *The Routledge handbook of language and intercultural communication* (pp. 85–97). Routledge.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/1.1.1

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Dong, X. (2009). *Requests in academic settings in English, Russian and Chinese* [Doctoral dissertation, The Ohio State University]. ProQuest Dissertations & Theses Global.

Ellis, R. (2003). *The study of second language acquisition*. Oxford University Press.

Han, S. (2005). *The interlanguage pragmatic development of the speech act of requests by Korean non-native speakers of English in an ESL setting* (Publication No. AAI3165689) [Doctoral dissertation, University of Pennsylvania]. ProQuest Dissertations & Theses Global.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin Books.

Ishihara, N. (2010). Teaching pragmatics in the classroom. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 101–124). Cambridge University Press.

Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Pearson Education.

Kachru, B. B. (1982). *The other tongue: English across cultures*. University of Illinois Press.

Kádár, D. Z., & Haugh, M. (2013). *Understanding politeness*. Cambridge University Press.

Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition, 13*(2), 215–247.

Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Blackwell.

Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition, 18*(2), 149–169.

Kecskés, I. (2014). *Intercultural pragmatics*. Oxford University Press.

Kecskés, I. (2016). Cross-cultural and intercultural pragmatics. In Y. Huang (Ed.), *The Oxford handbook of pragmatics* (pp. 1–16). Oxford University Press.

Leech, G. N. (1983). *Principles of pragmatics*. Longman.

Liddicoat, A. J. (2022). Intercultural mediation in language learning. In I. Kecskes (Ed.), *The Cambridge handbook of intercultural pragmatics* (pp. 815–835). Cambridge University Press. https://doi.org/10.1017/9781108884303.032

McConachy, T. (2018). *Developing intercultural perspectives on language use: Exploring pragmatics and culture in foreign language learning*. Multilingual Matters.

McConachy, T. (2022). Pragmatic awareness in intercultural language learning. In I. Kecskes (Ed.), *The Cambridge handbook of intercultural pragmatics* (pp. 766–787). Cambridge University Press.

Ogiermann, E. (2009). Politeness and indirectness across cultures: A comparison of English, German, Polish and Russian requests. *Journal of Politeness Research, 5*(2), 189–216.

Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System, 33*(3), 385–399.

Schauer, G. A. (2006). Pragmatic awareness in ESL and EFL contexts: Contrast and development. *Language Learning, 56*(2), 269–318.

Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge University Press.

Spencer-Oatey, H. (2008). Face, (im)politeness and rapport. In H. Spencer-Oatey (Ed.), *Culturally speaking: Culture, communication and politeness theory* (2nd ed., pp. 11–47). Continuum.

Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching, 48*(1), 1–50.

Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford University Press.

Takahashi, S. (2010). Instructional effects on request strategies in Japanese EFL learners. *Language Learning, 60*(1), 1–33.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*(2), 91–112. https://doi.org/10.1093/applin/4.2.91

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Bionote

Anh Nguyen-Hoang-Vuong received a B.A. in English Language Teaching from Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Vietnam, in 2009, and an M.Ed. in TESOL from the University of Southern Queensland, Australia, in 2013. She has completed various professional development programs in TESOL, action research, language testing and assessment, and intercultural communicative competence. She has been a lecturer at HUFLIT since 2010 and is currently a Ph.D. candidate in TESOL at Ho Chi Minh

City Open University, Vietnam. Her research interests include language and culture, teaching methodology, sociolinguistics, and nonverbal communication.

Appendices

Appendix A: Questionnaire on Students' Perceptions of Role-Play and Pragmatic Development

Part 1: Please indicate your level of agreement with each statement by selecting one option from 1 to 5.

1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree

No.	Statements	1	2	3	4	5
1	Role-play activities helped me learn how to make appropriate requests in different situations.	☐	☐	☐	☐	☐
2	Role-play activities improved my ability to use English in real-life communication.	☐	☐	☐	☐	☐
3	Role-play activities made the speaking lessons more engaging and interesting.	☐	☐	☐	☐	☐
4	I can distinguish between direct and indirect ways of making requests.	☐	☐	☐	☐	☐
5	I am able to adjust my language depending on the social context (e.g., teacher vs. friend).	☐	☐	☐	☐	☐
6	I understand how power relations (e.g., student–teacher) affect the way requests are made.	☐	☐	☐	☐	☐
7	I feel more aware of cultural differences when making requests in English.	☐	☐	☐	☐	☐
8	I am able to use strategies (e.g., hedging, softening) to make my requests more polite.	☐	☐	☐	☐	☐
9	I am aware that inappropriate requests may cause misunderstanding in intercultural communication.	☐	☐	☐	☐	☐
10	I realize that a request might be perceived differently by a native speaker compared to a Vietnamese speaker.	☐	☐	☐	☐	☐
11	I feel more confident when making requests in English.	☐	☐	☐	☐	☐
12	I feel less anxious when speaking English in role-play activities.	☐	☐	☐	☐	☐
13	I would like to continue using role-play activities in future English classes.	☐	☐	☐	☐	☐

Part 2: Please choose the option that best reflects your current practice.

14. When making a request, what do you usually focus on most?

A. Grammar accuracy only

B. Both grammar and appropriateness in context

C. Politeness strategies only

D. Getting the meaning across regardless of forms

15. How do you usually decide which request form to use?

A. I use the same form in all situations

B. I depend on the situation and the listener (Power, Distance, Imposition)

C. I translate directly from Vietnamese

D. I follow my personal habits and natural reflex

16. What difficulty do you face most when making requests in English?

A. Choosing correct grammar

B. Choosing appropriate level of politeness

C. Lack of vocabulary

D. Feeling anxious or shy to ask

17. Your requests are influenced by:

A. Vietnamese language habits (L1 Transfer)

B. Textbook-based learning only

C. A combination of classroom learning and real-life awareness

D. Media influence (movies, social media)

18. Compared to before, your English communication now is:

A. Still mainly textbook-based (EFL-oriented)

B. More practical and flexible (ESL-oriented)

C. Similar to real-life communication

D. No significant change

19. What kind of feedback do you receive from the teacher?

A. Only grammar correction

B. Feedback on appropriateness (politeness, context)

C. Feedback on both language and culture

D. General feedback on participation only

20. When requesting someone with higher status (e.g., a boss or teacher), you usually:

A. Use direct requests to save time

B. Use indirect and polite requests

C. Use very long explanations before the request

D. Feel hesitant and tend to avoid the request

21. When your request is misunderstood, you usually:

A. Repeat the same sentence or stop communicating

B. Change the wording or strategy (negotiation of meaning)

C. Switch to Vietnamese

D. Use gestures or ask for help from others

22. Which factor influences your choice of request strategy the most?

A. Relationship with the listener and social distance

B. The importance or difficulty of the favor

C. English grammar proficiency

D. Personal mood and confidence

23. Where do you mostly practice making requests in English?

A. Grammar exercises in textbooks

B. Role-play activities and simulations in class

C. In real-life communication outside class

D. Through online platforms/social media

24. Which activity helps you understand social variables (Power, Distance) best?

A. Teacher's theoretical explanation

B. Reading textbook dialogues

C. Participating in role-play scenarios

D. Observing native speakers (videos, movies)

25. After the course, how do you evaluate your request performance?

A. No significant change

B. Can perform requests but still feel culturally awkward

C. Can perform requests naturally and appropriately

D. Need much more real-life practice

Appendix B: Written Discourse Completion Test (WDCT)

Please read each situation carefully and imagine that you are in the described context. Write what you would naturally say to make an appropriate request in English.

When responding, please consider:
- *The relationship between you and the listener.*
- *The social context of the interaction.*
- *The level of politeness required.*

No.	Situation	Responses
1	**Equal Power, Informal Context** You are talking to your close friend in class. You forgot to bring your notebook today. What would you say to ask your friend to lend you their notebook?	
2	**Higher Power, Formal Context** You need to ask your lecturer for an extension on an assignment because you have been sick. What would you say to your lecturer?	
3	**Lower Power Context** You are the leader of a group project, and one of your group members has not completed their assigned task. What would you say to ask them to finish their work?	
4	**Stranger, Service Encounter** You are in a café, and the music is too loud. You would like a staff member to lower the volume. What would you say?	
5	**High-Imposition Request** You would like to ask a classmate, whom you do not know very well, to help you with a difficult assignment this weekend. What would you say?	
6	**Intercultural Context** You are studying abroad. You would like to ask a classmate from another cultural background to explain something that you do not understand. What would you say?	

Appendix C: List of Abbreviations

CCSARP: *Cross-Cultural Study of Speech Act Realization Patterns*

EFL: *English as a Foreign Language*

ESL: *English as a Second Language*

HUFLIT: *Ho Chi Minh City University of Foreign Languages and Information Technology*

L1: *First Language*

WDCT: *Written Discourse Completion Task*

#50 Using Short-Form Videos (TikTok/Reels) to Enhance Real-Life Listening and Speaking Skills for A2 EFL Students: A Pilot Study at Hùng Vương University of Ho Chi Minh City

Ông Hồng Thanh[1]

Abstract

A persistent challenge in Vietnamese EFL education was the gap between classroom learning and real-world communication. Many non-English major students performed well on grammar tests but struggled to understand natural speech or engage in spontaneous conversation. Meanwhile, these students frequently consumed short-form video content on platforms like TikTok and Reels, much of which featured authentic, unscripted English. This study explored whether this daily habit could be leveraged as a pedagogical tool.

This pilot study investigated the use of curated short-form videos to improve listening comprehension and speaking confidence among second-year, non-English major students at Hùng Vương University (CEFR A2 level). Over eight weeks, approximately 30 students engaged with videos depicting real-life scenarios—ordering food, casual conversations, and everyday transactions. Classroom activities included shadowing for pronunciation, learning colloquial expressions, role-plays, and student-created videos to reduce speaking anxiety.

Using a pretest-posttest design, the study combined listening and speaking assessments, questionnaires on communication confidence, and focus group interviews. Expected outcomes included improved ability to decode natural speech and increased willingness to communicate. Positive findings would have offered Vietnamese EFL teachers a low-cost, engaging method to enhance communicative competence using familiar digital platforms. This pilot provided foundational insights for future research and curriculum development at Hùng Vương University and beyond.

Keywords: short-form videos, TikTok, listening skills, speaking skills, A2 level, EFL, authentic materials, willingness to communicate

1. INTRODUCTION

1.1. The Teaching Reality and the Speaking-Listening Gap

A persistent challenge in Vietnamese EFL education is the gap between what students learn in the classroom and what they need for real-world communication. Many non-English major students can perform well on grammar tests but struggle with authentic listening or spontaneous interaction. At Hung Vuong University of Ho Chi Minh City (DHV), this problem is particularly evident among second-year, non-English major students at the CEFR A2 level. Research by Dang et al. (2020) found that 76% of first-year non-English major students at HVU come from rural areas in Northern Vietnam. Although these students have studied English for 7–10 years, they "only focused on English grammar for the final tests and had no practice in speaking and listening skills" (Dang et al., 2020, p. 91). As a result, they scored only 10–15 out of 25 on the KET listening and speaking placement test, confirming their A2 proficiency level.

[1] *Faculty of Languages, Hung Vương University of Ho Chi Minh City (DHV), Vietnam; Email: thanhoh@dhv.edu.vn*

This phenomenon is not unique to HVU. Nationwide, Vietnamese learners of English often exhibit high anxiety and low willingness to communicate (WTC) when speaking English. Cong-Lem and Nguyen, 2024 found that willingness to communicate correlates positively with speaking proficiency but negatively with speaking test anxiety, and importantly, that students are more willing to communicate outside the classroom than inside it. This finding suggests that authentic, out-of-school materials—such as short-form videos on TikTok and Reels—could serve as a powerful bridge to increase learners' WTC.

The root of this problem lies in the disconnect between textbook English and natural speech. Textbooks present slow, carefully enunciated dialogues that lack the phonological features of real conversation—reductions ("gonna," "wanna"), linking, elision, and fillers ("um," "like") (McCarthy & Carter, 1995). When A2 learners encounter authentic speech, they often fail to recognize familiar words because the acoustic signal does not match their mental representation (Field, 2008). This input gap leads to frustration, low motivation, and avoidance of speaking opportunities.

1.2. The Rise of Short-Form Videos as a Potential Solution

Paradoxically, while students avoid authentic English materials, they spend considerable time on short-form video platforms. Recent data show that TikTok has become a global phenomenon, with users spending an average of 95 minutes per day on the app (Duarte, 2026, cited in Do & Huynh, 2026). In Vietnam, TikTok's popularity is exceptionally high, making it a relevant context for educational research (Le, 2023).

Several recent Vietnamese studies have explored TikTok's potential for language learning. Do and Huynh (2026) surveyed 105 university students across Vietnam and found that TikTok significantly improves listening (75.24% approval) and speaking (67.62% approval) skills, with strong correlations to increased confidence (0.729) and motivation (0.498) for English use beyond the classroom. Similarly, Ngo (2024) reported that Vietnamese college students perceive TikTok as an effective tool for independent English speaking practice, citing its authentic content, repetition, and low-anxiety environment. Le (2023), through semi-structured interviews with Vietnamese teachers and students in Hanoi, found that TikTok is viewed as a "fun and engaging" and "effective way to engage and motivate learners," though the study also noted challenges such as lack of context, informal language, and addictive characteristics.

These studies align with microlearning theory, which suggests that short, focused content segments (15–60 seconds) reduce cognitive load and enhance retention (Gabrielli et al., 2017; Hug, 2005, both cited in Do & Huynh, 2026). Short-form videos fit this model perfectly: they are brief, repetitive, visually supportive, and highly engaging.

1.3. The Research Gap

Despite these promising findings, significant gaps remain. First, most existing studies on short-form videos in Vietnam have focused on students' perceptions rather than measuring actual learning outcomes through pre-/post-test designs. While Do and Huynh (2026) and Ngo (2024) provide valuable survey data, neither employed an experimental or quasi-experimental design to measure skill improvement. Second, no study has specifically targeted A2-level non-English major students at a university like HVU, where students have limited prior exposure to authentic English and low speaking confidence. Third, while studies like Ta and Tran (2024) at Tay Do University and Vu (2024) at Huynh Khuong Ninh Secondary School have shown positive impacts of TikTok on speaking skills, they did not integrate structured techniques such as shadowing—a well-established method for improving pronunciation and listening through

immediate repetition (Hamada, 2015). Finally, there is a notable lack of research combining short-form video consumption with structured, in-class speaking activities (shadowing, role-play, and student-created videos) for A2 learners in the Vietnamese context. This pilot study aims to fill these gaps by providing preliminary evidence from a structured, eight-week intervention at Hung Vuong University.

1.4. Research Questions

To address the gaps identified above, this study seeks to answer the following three research questions:

To what extent does an eight-week intervention using curated short-form videos (TikTok/Reels) affect second-year, non-English major A2 students' ability to comprehend authentic spoken English, as measured by a customized listening pre-test and post-test?

How does participation in the intervention influence students' self-reported confidence, speaking anxiety, and willingness to communicate (WTC) when using English for everyday topics, as measured by pre- and post-intervention questionnaires?

What are students' perceptions of the intervention's relevance, engagement, and usefulness as a learning tool for real-life communication, as revealed through semi-structured focus group interviews?

1.5. Significance of the Study

This pilot study, despite its modest scale, carries meaningful implications for three main groups.

For classroom teachers, this study offers a practical, low-cost, and engaging pedagogical model. Teachers do not need expensive software or specialized training—only a smartphone, an internet connection, and a willingness to explore platforms students already use daily. The structured four-stage lesson plan (warm-up, video viewing with comprehension, shadowing for pronunciation, and interactive role-play) can be adapted to various teaching contexts, from high schools to language centers. As Vu (2024) demonstrated at a secondary school in Vietnam, TikTok-based activities significantly improved speaking scores and student confidence, with most students holding a positive view of TikTok as an educational tool.

For curriculum developers and course designers, this study provides evidence to reconsider the role of authentic digital materials in official syllabi. Currently, most general English programs at Vietnamese universities rely almost exclusively on commercial textbooks. If this pilot yields positive results, it could encourage curriculum designers to formally integrate short-form video activities as a regular component of listening and speaking courses—not merely as occasional "fun Friday" activities. As Do and Huynh (2026) argue, integrating TikTok into the EFL curriculum can enhance motivation and communication skills, offering a "dynamic supplement to traditional classroom instruction."

For future researchers, this study serves as a foundation. As a pilot, its purpose is not to make grand claims but to test the feasibility of the intervention, refine data collection instruments, and generate preliminary data. Positive results would justify a larger-scale study with a control group, longer duration, and multiple institutions. Negative results would help identify what does not work for A2 learners, saving future researchers time and effort. Additionally, this study contributes to the growing international conversation about harnessing social media for meaningful language learning while providing much-needed context-specific data from a Vietnamese university in a non-metropolitan setting.

2. LITERATURE REVIEW

2.1. The Challenge of Authentic Listening and Spoken Interaction for A2 Learners

One of the most robust findings in second language acquisition (SLA) research is the significant gap between the careful, scripted speech found in textbooks and the rapid, reduced speech of everyday conversation (McCarthy & Carter, 1995). Native speakers produce speech characterized by phonological reductions (e.g., "going to" → "gonna"), elision, linking, and fillers (Field, 2008). For A2 learners, these features pose substantial challenges. In the Vietnamese context, Dang et al. (2020) found that DHV students had almost no prior practice in listening and speaking. Tran, Dang, and Nguyen (2024) identified pronunciation, intonation, anxiety, and lack of exposure to authentic materials as key barriers to listening comprehension.

2.2. Authentic Materials in EFL Teaching

The use of authentic materials has been advocated since the rise of communicative language teaching (Nunan, 1989). However, a long-standing concern is that authentic materials are often too difficult for lower-proficiency learners (Guariento & Morley, 2001). Short-form videos offer a solution due to their brevity (15–60 seconds), which reduces cognitive load and allows for repeated viewing (Do & Huynh, 2026).

2.3. Short-Form Videos in EFL Education

Global studies have shown positive results. Xiuwen and Razali (2021) found that TikTok improved Malaysian EFL students' pronunciation and motivation. Zaitun, Hadi, and Indriani (2021) reported reduced speaking anxiety among Indonesian students. Efeoğlu (2023) found that eight weeks of shadowing with short videos significantly improved Turkish EFL learners' pronunciation.

Within Vietnam, Do and Huynh (2026) found that 75.24% of participants agreed TikTok improved listening skills, and 67.62% reported improved speaking skills. Ta and Tran (2024) at Tay Do University found TikTok-based activities helped improve fluency and confidence. Vu (2024) reported significant improvements in speaking scores at a secondary school.

2.4. Shadowing for Pronunciation and Listening Development

Shadowing—listening to a short segment and immediately repeating it aloud—improves listening comprehension, pronunciation, and automaticity (Hamada, 2015). In Vietnam, Phan, Ly, and Nguyen (2024) found that video-based shadowing significantly improved learners' pronunciation, particularly intonation.

2.5. Willingness to Communicate and Speaking Anxiety

WTC refers to a learner's readiness to engage in communication (MacIntyre et al., 1998). Anxiety is a strong negative predictor of WTC. In Vietnam, Hoang (2022) found that communication apprehension was the most common type of anxiety, with a negative correlation between anxiety and WTC. Ngo and Nguyen (2024) found that students were more willing to communicate outside the classroom than inside it, suggesting authentic materials may lower anxiety.

2.6. Conceptual Framework

This study is guided by three theoretical pillars: (1) authentic input (Krashen, 1982), (2) structured output through shadowing (Hamada, 2015; Swain, 1985), and (3) lowered affective filter leading to increased WTC (Krashen, 1982).

3. METHODOLOGY

3.1. Research Design

This study employed a single-group pretest-posttest design, appropriate for a pilot study (Creswell & Creswell, 2018).

3.2. Participants

Participants were second-year, non-English major students at Hung Vuong University of Ho Chi Minh City (DHV). After screening for incomplete data, 22 students (10 male, 12 female) had complete data for both listening and speaking tests and were included in the final analysis. All participants were assessed at CEFR A2 level.

3.3. The Intervention: An Eight-Week Program

The intervention consisted of eight weekly 90-minute sessions. Each week focused on a specific theme: ordering food, making small talk, giving compliments, asking for directions, expressing opinions, handling service encounters, making plans, and reacting to news.

Each 90-minute session followed a 5-stage structure:

Stage	Activity	Duration
1	Warm-up (discussing the theme)	10 min
2	Video viewing & comprehension (2-3 videos)	20 min
3	Shadowing & pronunciation practice	20 min
4	Interactive role-play (unscripted scenarios)	30 min
5	Student-created video assignment	10 min + homework

3.4. Data Collection Instruments

3.4.1. Listening and Speaking Tests

A customized listening test (12 questions, maximum score 10) and a speaking test (maximum score 10) were administered before (pretest) and after (posttest) the intervention.

3.4.2. Questionnaire

A 5-point Likert scale questionnaire measured self-reported confidence, anxiety, and willingness to communicate (WTC) before and after the intervention.

3.4.3. Focus Group Interviews

Semi-structured interviews were conducted with 5 students representing different proficiency levels (weak, average, good). Interviews were conducted in Vietnamese, transcribed, and translated into English.

3.5. Data Analysis

Paired-samples t-tests were used to compare pre- and post-test scores. Effect sizes (Cohen's d) were calculated. Interview data were analyzed using thematic analysis (Braun & Clarke, 2006).

4. FINDINGS / RESULTS

4.1. Participants

A total of 30 second-year non-English major students at Hung Vuong University of Ho Chi Minh City (DHV) initially participated in the study. After screening for incomplete data (students with zero scores on either the pre-test or post-test, indicating absence on testing days), **22 students** had complete data for both listening and speaking tests. These 22 students were included in the final analysis.

Among the 22 participants:

Gender: 10 male (45.5%), 12 female (54.5%)

Age range: 19 to 21 years (Mean = 19.8, SD = 0.65)

Prior English study: All had studied English for 7-10 years in primary and secondary school

Social media usage: All 22 students reported using TikTok or Instagram Reels daily, with an average daily usage of 112 minutes (approximately 1 hour 52 minutes). Only 6 students (27.3%) reported that they had previously used TikTok content for English learning purposes before the intervention.

4.2. Results for Research Question 1: Listening Comprehension

Research Question 1 asked: *To what extent does an eight-week intervention using curated short-form videos affect A2 students' ability to comprehend authentic spoken English?*

A customized listening test (maximum score = 10) was administered before (pretest) and after (posttest) the eight-week intervention. Descriptive statistics are presented in Table 1.

Table 1

Descriptive Statistics for Listening Test Scores (n = 22)

Test	Mean	Standard Deviation (SD)	Minimum	Maximum
Pre-test	6.48	1.66	2.8	10.0
Post-test	7.63	2.13	0	9.1

As shown in Table 1, the mean listening score increased from 6.48 (SD = 1.66) on the pre-test to 7.63 (SD = 2.13) on the post-test, representing an average gain of 1.15 points (approximately 17.7% improvement).

Figure 1

Illustrates the individual score changes from pre-test to post-test:

Score change category	Number of students	Percentage
Improved (post-test > pre-test)	16	72.7%
No change (post-test = pre-test)	1	4.5%
Declined (post-test < pre-test)	5	22.7%

The majority of students (72.7%) showed improvement in their listening comprehension after the eight-week intervention.

A paired-samples t-test was conducted to determine whether the difference between pre-test and post-test scores was statistically significant. The results showed that post-test scores (M = 7.63, SD = 2.13) were significantly higher than pre-test scores (M = 6.48, SD = 1.66), with t(21) = 3.21, p = .004. The effect size was moderate to large (Cohen's d = 0.68), indicating that the intervention had a meaningful practical impact on students' listening comprehension.

Analysis by proficiency level: To gain a deeper understanding, students were divided into three subgroups based on their pre-test listening scores:

Proficiency level	n	Pre-test mean	Post-test mean	Gain	p-value
Low (pre-test < 5.0)	4	3.9	7.8	+3.9	.008
Medium (5.0 - 7.9)	12	6.1	7.5	+1.4	.031
High (≥ 8.0)	6	8.9	7.5	-1.4	.082

Interestingly, the lowest proficiency students showed the largest gains (from 3.9 to 7.8, an increase of 3.9 points, p = .008). This suggests that the intervention was particularly beneficial for students who started with the weakest listening skills. In contrast, high-proficiency students showed a slight (non-significant) decline, possibly due to ceiling effects or because the video content was not challenging enough for them.

4.3. Results for Research Question 2: Speaking Performance

Research Question 2 asked: How does participation in the intervention influence students' speaking performance?

A speaking test (maximum score = 10) was administered before and after the intervention. The speaking test assessed fluency, pronunciation, vocabulary, and comprehensibility using the rubric described in Appendix C.

Table 2: Descriptive Statistics for Speaking Test Scores (n = 22)

Test	Mean	Standard Deviation (SD)	Minimum	Maximum
Pre-test	7.25	1.89	2.8	10.0
Post-test	7.68	1.72	0	9.1

As shown in Table 2, the mean speaking score increased from 7.25 (SD = 1.89) on the pre-test to 7.68 (SD = 1.72) on the post-test, representing an average gain of 0.43 points (approximately 5.9% improvement).

Figure 2

Illustrates the individual score changes for speaking:

Score change category	Number of students	Percentage
Improved (post-test > pre-test)	13	59.1%

Score change category	Number of students	Percentage
No change (post-test = pre-test)	2	9.1%
Declined (post-test < pre-test)	7	31.8%

More than half of the students (59.1%) showed improvement in their speaking performance, though the improvement was smaller than for listening.

A paired-samples t-test was conducted. The results showed that post-test scores (M = 7.68, SD = 1.72) were higher than pre-test scores (M = 7.25, SD = 1.89), but the difference was not statistically significant, with t(21) = 1.45, p = .162. The effect size was small (Cohen's d = 0.31).

Analysis by proficiency level:

Proficiency level	n	Pre-test mean	Post-test mean	Gain	p-value
Low (pre-test < 5.0)	3	4.0	7.3	+3.3	.042
Medium (5.0 - 7.9)	10	6.8	7.5	+0.7	.187
High (≥ 8.0)	9	8.9	8.1	-0.8	.095

Similar to listening, the lowest proficiency students showed the largest speaking gains (from 4.0 to 7.3, an increase of 3.3 points, p = .042). However, the small number of students in the low-proficiency group (n=3) means this finding should be interpreted with caution.

4.4. Comparison of Listening vs. Speaking Gains

Table 3:

Comparison of Listening and Speaking Improvements (n=22)

Measure	Pre-test Mean	Post-test Mean	Mean Gain	t-value	p-value	Cohen's d
Listening	6.48	7.63	+1.15	3.21	.004	0.68
Speaking	7.25	7.68	+0.43	1.45	.162	0.31

The data clearly show that the intervention had a stronger and statistically significant effect on listening comprehension than on speaking performance. This finding is discussed further in Section 5.

4.5. Results for Research Question 3: Confidence, Anxiety, and WTC

Research Question 3 asked: How does participation in the intervention influence students' self-reported confidence, speaking anxiety, and willingness to communicate (WTC)?

A 20-item questionnaire using a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) was administered before and after the intervention. The questionnaire measured three constructs: (a) speaking confidence (8 items), (b) speaking anxiety (7 items, reverse-coded), and (c) willingness to communicate (5 items).

Table 4:

Descriptive Statistics for Questionnaire Scales (n = 22)

Construct	Pre-test Mean	Post-test Mean	Mean Change	Interpretation
Confidence (1-5)	2.0	4.0	+2.0	Large increase
Anxiety (1-5)	3.5	3.0	-0.5	Small decrease
WTC(1-5)	4.0	4.5	+0.5	Small increase

Confidence: Students reported a substantial increase in speaking confidence, from 2.0 to 4.0 out of 5. This represents a 100% improvement. This finding is consistent with Do and Huynh (2026), who found a strong correlation (r = 0.729) between TikTok use and increased confidence.

Anxiety: Speaking anxiety decreased slightly from 3.5 to 3.0. While the decrease was modest, it is notable that anxiety did not increase despite the introduction of student-created video assignments, which some students initially found stressful.

Willingness to communicate (WTC): WTC increased slightly from 4.0 to 4.5. This suggests that students became more willing to seek out and engage in English communication opportunities after the intervention.

Individual item analysis: Among the confidence items, the largest gains were observed for:

"I feel confident speaking English in real-life situations": Pre = 2.1, Post = 4.2 (+2.1)

"I can order food in English without feeling nervous": Pre = 1.9, Post = 4.3 (+2.4)

"I can understand native speakers in short videos": Pre = 2.3, Post = 4.1 (+1.8)

Among the anxiety items, the largest reductions were observed for:

"I feel anxious when I have to speak English without preparation": Pre = 4.0, Post = 2.8 (-1.2)

"I feel nervous recording myself speaking English": Pre = 4.2, Post = 3.2 (-1.0)

4.6. Qualitative Findings: Student Perceptions

To answer Research Question 4 (students' perceptions), five students representing different proficiency levels (weak, average, good) participated in semi-structured focus group interviews. Thematic analysis (Braun & Clarke, 2006) revealed five major themes.

Theme 1: Authenticity – "It helps me understand real-life language use"

Four out of five students mentioned that the videos helped them understand how English is actually used in daily life.

"These videos help me understand how to use words in real life." (Chế Thanh Tuyền, female, average)

"It helps me understand natural English communication better, especially slang and intonation." (Khưu Tuấn Kiệt, male, good)

"The videos show me how native speakers actually talk, not just textbook English." (Hà Hiếu Kỳ, male, good)

Theme 2: Shadowing as the most useful technique

Three students specifically identified shadowing as the most helpful activity.

"I like shadowing the most because it helps me imitate native speaker pronunciation." (Chế Thanh Tuyền, female, average)

"Shadowing helps me learn the rhythm of English. My mouth moves differently now." (Huỳnh Gia Mẫn, female, good)

Theme 3: Fast speech as a persistent challenge

All students mentioned that the speed of natural speech was challenging, particularly in the first few weeks.

"Sometimes TikTok videos are too fast, so I have difficulty keeping up." (Chế Thanh Tuyền, female, average)

"I don't like it much because there are many difficult words and they speak too fast. But watching many times helps a little." (Nguyễn Anh Thư, female, weak)

"At first, I couldn't understand anything. But after a few weeks, my ears got used to the speed." (Hà Hiếu Kỳ, male, good)

Theme 4: Anxiety about creating videos (initially)

Four out of five students reported initial anxiety about recording themselves speaking English.

"I am a bit shy when recording videos because I'm afraid of saying something wrong." (Chế Thanh Tuyền, female, average)

"I feel pressured when recording videos because I have to speak fluently and confidently." (Hà Hiếu Kỳ, male, good)

"I am very shy about recording videos because I am not confident in my English." (Nguyễn Anh Thư, female, weak)

Theme 5: Increased confidence over time

Despite initial anxiety, students reported that their confidence grew with repeated practice.

"At first, I was worried because I thought others would judge me. But after doing it many times, I feel more confident." (Huỳnh Gia Mẫn, female, good)

"I am not worried about recording videos because I see it as an opportunity to practice speaking. The more I do it, the easier it gets." (Khưu Tuấn Kiệt, male, good)

"The first video took me 20 tries. Now I can do it in one or two takes." (Chế Thanh Tuyền, female, average)

Theme 6: Recommendation for future students

All five students said they would recommend the course to future students, with some suggestions for improvement.

"Yes, I would recommend it. But maybe make some videos slower or provide transcripts." (Nguyễn Anh Thư, female, weak)

"Definitely yes. This should be a required course for all A2 students." (Khưu Tuấn Kiệt, male, good)

4.7. Summary of Key Findings

Research Question	Key Finding	Statistical Significance
RQ1 (Listening)	Significant improvement (+1.15 points)	p = .004, d = 0.68
RQ1 (Listening by level)	Lowest proficiency students gained most (+3.9)	p = .008
RQ2 (Speaking)	Improvement but not significant (+0.43)	p = .162, d = 0.31
RQ2 (Speaking by level)	Lowest proficiency students gained most (+3.3)	p = .042
RQ3 (Confidence)	Large increase (2.0 → 4.0)	Practical significance
RQ3 (Anxiety)	Small decrease (3.5 → 3.0)	Practical significance
RQ3 (WTC)	Small increase (4.0 → 4.5)	Practical significance
RQ4 (Perceptions)	Positive overall; shadowing most useful; fast speech challenging; video Anxiety decreases with practice	Thematic Analysis

5. DISCUSSION

5.1. Interpretation of Findings

The results of this pilot study provide several important insights into the use of short-form videos for A2 EFL learners.

Listening comprehension improved significantly. The finding that listening scores increased from 6.48 to 7.63 (p = .004, d = 0.68) aligns with the input hypothesis (Krashen, 1982). The short-form videos provided compelling, comprehensible input just above the students' current level (i+1). The visual nature of the videos, combined with their brevity and repetition, likely made authentic input accessible to A2 learners who would otherwise find unscripted speech overwhelming.

Speaking improvement was not statistically significant. While speaking scores increased from 7.25 to 7.68, this change was not significant (p = .162). There are several

possible explanations. First, eight weeks may be insufficient time for measurable speaking gains, as speaking requires not only input but also extensive output practice and neurological changes in motor planning for pronunciation. Second, the pretest speaking scores were already relatively high (7.25), leaving less room for improvement. Third, speaking is a more complex skill than listening, requiring real-time processing, pronunciation, grammar, and pragmatics simultaneously.

Confidence increased substantially. The increase in self-reported confidence from 2.0 to 4.0 (out of 5) is striking. This aligns with the affective filter hypothesis (Krashen, 1982). The familiar, low-stakes, entertaining nature of TikTok likely lowered students' anxiety, allowing them to engage more willingly with English. This finding is consistent with Do and Huynh (2026), who found a strong correlation ($r = 0.729$) between TikTok use and increased confidence.

Anxiety decreased slightly. The small decrease in anxiety ($3.5 \rightarrow 3.0$) is less dramatic than expected, but the interview data provide an explanation. Students reported initial anxiety about creating their own videos. As one student said, "I am a bit shy when recording videos because I'm afraid of saying something wrong" (Tuyền). However, students also reported that anxiety decreased with repeated practice: "After doing it many times, I feel more confident" (Mẫn).

WTC increased modestly. The increase in willingness to communicate ($4.0 \rightarrow 4.5$) suggests that students became more willing to speak English after the intervention. This supports Cong-Lem and Nguyen's (2024) finding that students are more willing to communicate outside the classroom.

5.2. Comparison with Previous Research

These findings are consistent with previous studies. Efeoğlu (2023) found that eight weeks of shadowing with short videos improved Turkish EFL learners' pronunciation. Similarly, Phan, Ly, and Nguyen (2024) found that Vietnamese EFL learners held positive attitudes toward video-based shadowing. Do and Huynh (2026) reported that 75.24% of Vietnamese students perceived TikTok as improving their listening skills, which aligns with our quantitative finding of significant listening gains.

However, our finding of non-significant speaking improvement differs from some previous studies. Zaitun et al. (2021) reported reduced speaking anxiety and improved speaking skills among Indonesian students. This discrepancy may be due to differences in proficiency levels (our students were A2), intervention duration (eight weeks may be too short), or assessment methods.

5.3. Pedagogical Implications

Based on these findings, several practical recommendations can be offered:

Short-form videos are effective for listening development. Teachers should integrate curated TikTok/Reels content regularly, not as occasional activities.

Shadowing should be explicitly taught. Students need instruction on how to mimic intonation, rhythm, and reductions.

Student-created videos require scaffolding. Teachers should provide clear instructions, allow multiple takes, and create a supportive, non-judgmental environment. As the interview data show, initial anxiety is common but decreases with practice.

Fast speech remains a challenge. Teachers should provide transcripts, slow down videos when possible, and repeat key segments.

Longer interventions may be needed for speaking gains. Eight weeks was sufficient for listening but not for speaking. Future implementations should consider 12-16 weeks.

5.4. Limitations

Several limitations must be acknowledged:

No control group. Without a comparison group, causality cannot be definitively established.

Small sample size (n = 22). Results may not generalize to other contexts.

Short duration. Eight weeks may be insufficient for speaking gains.

Single institution. DHV is a specific context; results may differ elsewhere.

Researcher as instructor. Potential bias; future studies should use blind assessors.

Speaking assessment. The speaking test may not have captured all aspects of oral proficiency.

6. CONCLUSION

6.1. Summary of Findings

This pilot study investigated the effects of an eight-week intervention using curated short-form videos (TikTok/Reels) on the listening comprehension, speaking performance, confidence, anxiety, and willingness to communicate of 22 A2-level, second-year non-English major students at Hung Vuong University of Ho Chi Minh City (DHV).

The key findings are:

Finding	Result
Listening comprehension	Significant improvement ($p = .004$, $d = 0.68$)
Speaking performance	Improvement but not significant ($p = .162$)
Self-reported confidence	Large increase (2.0 → 4.0 out of 5)
Self-reported anxiety	Small decrease (3.5 → 3.0)
Willingness to communicate	Small increase (4.0 → 4.5)
Student perceptions	Positive overall; shadowing most useful; fast speech challenging; initial video anxiety decreases with practice

6.2. Contributions

This study makes several contributions:

Provides actual quantitative data on listening and speaking gains from a structured short-form video intervention with A2 learners in Vietnam.

Demonstrates that short-form videos significantly improve listening comprehension for A2 EFL students ($p = .004$, $d = 0.68$).

Shows a substantial increase in confidence (100% improvement), supporting the affective filter hypothesis.

Offers a replicable pedagogical model (5-stage weekly lesson structure) for Vietnamese EFL teachers.

Provides rich qualitative data from student interviews, revealing both benefits (authenticity, shadowing) and challenges (fast speech, initial video anxiety).

6.3. Recommendations for Future Research

Based on this pilot study, future research should:

Include a control group to establish causality.

Extend the intervention duration to 12-16 weeks to assess speaking gains.

Increase sample size and include multiple institutions.

Use blind raters for speaking assessments.

Include follow-up testing (3-6 months post-intervention) to assess retention.

Investigate individual differences (e.g., does the intervention work better for students with high vs. low social media use?).

6.4. Final Remarks

This pilot study provides evidence that short-form videos—platforms that students already use daily—can be effectively harnessed to improve listening comprehension and confidence among A2 EFL learners in Vietnam. While speaking gains were not statistically significant, the positive trends and strong student perceptions suggest that longer interventions may yield speaking improvements.

The question is not whether TikTok and Reels will continue to dominate young people's attention—they will. The question is whether educators can creatively turn scrolling time into meaningful language learning time. This study suggests the answer is yes, at least for listening and confidence. With longer interventions and larger samples, speaking gains may follow.

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Cong-Lem, N., & Nguyen, T. T. H. (2024). *The interplay among L2 willingness to communicate, speaking test anxiety and speaking proficiency.* Bridges Monash University. https://doi.org/10.26180/25203293.v1

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dang, L. N. H., Vu, T. Q. D., Pham, T. T. H., & Nguyen, T. M. L. (2020). Some factors affecting English proficiency of the first year non-English majored students at Hung Vuong University. J*ournal of Science and Technology Hung Vuong University, 21*(4), 87-96.

Do, T. Y. L., & Huynh, N. K. (2026). Utilizing TikTok in improving English skills: A case study of Vietnamese university students. *International Journal of TESOL & Education, 6*(1), 37-56. https://doi.org/10.54855/ijte.26613

Efeoğlu, E. M. (2023). *The effects of shadowing technique on L2 learners' pronunciation skills* [Master's thesis, Necmettin Erbakan University].

Field, J. (2008). *Listening in the language classroom.* Cambridge University Press.

Guariento, W., & Morley, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal, 55*(4), 347-353. https://doi.org/10.1093/elt/55.4.347

Hamada, Y. (2015). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research, 20*(1), 35-52. https://doi.org/10.1177/1362168815597504

Hoang, T. K. O. (2022). An investigation into the influences of anxiety in non-English major tertiary learners' willingness to communicate orally in Vietnamese EFL classrooms. *Proceedings of the 4th Conference on Language Teaching and Learning.* AIJR Proceedings. https://doi.org/10.21467/proceedings.128

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Pergamon Press.

Le, A. T. (2023). *TikTok and its influence on the English competence of Vietnamese students* [Master's thesis, University of Alberta]. https://doi.org/10.7939/r3-b3cw-rw42

MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition, 23*(3), 369-388.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562.

McCarthy, M., & Carter, R. (1995). Spoken grammar: What is it and how can we teach it? *ELT Journal, 49*(3), 207-218. https://doi.org/10.1093/elt/49.3.207

Ngo, T. T. (2024). Vietnamese college students' perception towards using TikTok for independent English speaking practice. *Proceedings of the AsiaCALL International Conference, 6*, 112-122. https://doi.org/10.54855/paic.2468

Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge University Press.

Phan, V. T. T., Ly, K. L., & Nguyen, D. K. (2024). Effects of video-based shadowing on suprasegmental features: EFL learners' pronunciation performance and attitudes. *rEFLections, 31*(3), 896-924. https://doi.org/10.61508/refl.v31i3.275498

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.

Ta, H. M., & Tran, T. B. N. (2024). *An investigation into the impact of TikTok videos on improving students' speaking skills at a university in Can Tho City*. Tay Do University.

Tran, L. D. (2025). The efficacy of a translanguaging-based pedagogy in reducing speaking anxiety and improving grammatical accuracy in students at a university in Vietnam. *International Journal of Social Science and Human Research, 8*(7). https://doi.org/10.47191/ijsshr/v8-i7-08

Tran, T. T. A., Dang, N., & Nguyen, L. (2024). Factors affecting listening comprehension of English majored Vietnamese first-year students. *Advances in Social Sciences and Management, 2,* 53-68. https://doi.org/10.63002/assm.212.775

Vu, M. N. (2024). *Applying TikTok to improve students' speaking skill at Huynh Khuong Ninh Secondary School* [Master's thesis, Ba Ria – Vung Tau University].

Xiuwen, Z., & Razali, A. B. (2021). An overview of the utilization of TikTok to improve oral English communication competence among EFL undergraduate students. *Universal Journal of Educational Research, 9*(7), 1439-1445.

Zaitun, Z., Hadi, M. S., & Indriani, E. D. (2021). TikTok as a media to enhance the speaking skills of EFL students. *Jurnal Studi Guru dan Pembelajaran, 4*(2), 437-444.

AUTHOR BIOGRAPHY

Ong Hong Thanh is a lecturer at the Faculty of Languages, Hung Vương University of Ho Chi Minh City (DHV), Vietnam. She holds a Master's degree in TESOL from Edith Cowan University, Australia (2022). Her research interests include EFL listening and speaking pedagogy, technology-enhanced language learning, authentic materials development, and learner motivation. She has 11 years of experience teaching English to non-major students at the tertiary level. She is particularly interested in integrating social media platforms such as TikTok and Instagram Reels into English language teaching to bridge the gap between classroom learning and real-world communication. She can be contacted at thanhoh@dhv.edu.vn.

APPENDICES

Appendix A: Sample Lesson Plan (Week 1 – Ordering Food & Drinks)

Week 1 Theme: Ordering food and drinks

Duration: 90 minutes

Level: A2 (CEFR)

Materials: TikTok/Reels videos (pre-selected), projector/smartphone, handout, WhatsApp group

Lesson Objectives: By the end of the lesson, students will be able to:

Understand key phrases for ordering food/drinks from short videos

Shadow native speakers with improved pronunciation and rhythm

Role-play a café scenario using target language

Record a short video ordering their favorite drink

Lesson Procedure:

Stage	Time	Activity	Teacher's Role	Students' Role
1. Warm-up	10 min	**Discussion:** "What is your favorite drink? Where do you buy it? Have you ever ordered in English?"	Ask questions, write key vocabulary on board	Share answers in pairs, then with class
2. Vocabulary pre-teaching	5 min	Introduce key phrases: *"Can I get a...", "I'll have a...", "For here or to go?", "Could I have extra...", "That'll be..."*	Model pronunciation, explain meanings	Repeat after teacher, take notes
3. Video viewing (1st time)	5 min	Watch Video 1: "How to order coffee at Starbucks" (25 seconds)	Play video, ask: "What did the customer order?"	Watch for gist, answer question
4. Video viewing (2nd time)	5 min	Watch Video 1 again with pauses	Pause after each line, ask comprehension questions	Listen for details, answer
5. Video viewing (3rd time)	5 min	Watch Video 2: "Ordering boba tea" (20 seconds)	Play video, ask: "What size did she order? What topping?"	Watch, answer questions
6. Shadowing practice	15 min	Practice repeating key lines from both videos	Play line by line, model reductions (*"gonna", "wanna", "hafta"*), provide individual feedback	Repeat individually, then in pairs, then small groups

Stage	Time	Activity	Teacher's Role	Students' Role
7. Role-play preparation	10 min	Pairs receive scenario cards	Distribute cards, monitor, assist	Read scenario, prepare short dialogue (unscripted)
8. Role-play performance	15 min	Pairs perform for class	Observe, give brief positive feedback	Perform 30-45 second role-play
9. Video creation instruction	10 min	Explain homework: record 15-30 second video ordering favorite drink	Demonstrate with own video, explain uploading to WhatsApp group	Take notes, ask questions
10. Wrap-up & review	5 min	Review key phrases, announce homework deadline	Summarize, answer questions	Share one thing learned

Sample Role-Play Scenario Card (Week 1):

Scenario A (Customer): You are at a coffee shop. You want a medium hot latte with oat milk. The barista says they are out of oat milk.

Scenario A (Barista): You work at a coffee shop. A customer wants oat milk, but you only have regular milk and soy milk.

Students must improvise – no script allowed!

Homework Assignment (Week 1):

Record a **15-30 second video** ordering your favorite drink. You can use your smartphone. Upload to the class WhatsApp group before the next class.

Tips for students:

Watch the TikTok videos again before recording

Practice shadowing a few times first

You can re-record as many times as you want

Don't worry about perfection – just try your best!

Appendix B: Listening Test (Pretest and Posttest)

Test duration: 15 minutes

Total questions: 10

Maximum score: 10 points (1 point per correct answer)

Instructions (read to students):

"You will hear 5 short audio clips. Each clip is 15-30 seconds long. After each clip, you will answer 2 questions. Listen carefully. You may take notes. Each clip will be played twice."

Clip 1 (22 seconds) – Ordering coffee

Audio script (not shown to students): "Hi, can I get a medium latte with oat milk? And could you add an extra shot of espresso? Thanks."

Question	Answer choices	Correct answer
1. What does the person order?	A. Black coffee B. Latte C. Tea	**B**
2. What size do they want?	A. Small B. Medium C. Large	**B**

Clip 2 (18 seconds) – Making plans

Audio script: "Hey, are you free on Saturday? There's a new movie playing at the cinema. Want to catch the 7 PM show?"

Question	Answer choices	Correct answer
1. Where are they going?	A. Restaurant B. Cinema C. Park	**B**
2. What time will they meet?	A. 6:00 PM B. 6:30 PM C. 7:00 PM	**C**

Clip 3 (25 seconds) – Asking for directions

Audio script: "Excuse me, I'm looking for the nearest pharmacy. Is it far from here? I think I'm lost. Can you point me in the right direction?"

Question	Answer choices	Correct answer
1. Where is the person trying to go?	A. Museum B. Library C. Pharmacy	**C**
2. What is the person's problem?	A. They are lost B. They are late C. They are tired	**A**

Clip 4 (20 seconds) – Giving a compliment

Audio script: "Wow, I love your jacket! Where did you get it? That color really suits you. You have great style."

Question	Answer choices	Correct answer
1. What is the compliment about?	A. Hair B. Jacket C. Shoes	**B**
2. What does the speaker say about the color?	A. It's too bright B. It suits the person C. It's boring	**B**

Clip 5 (20 seconds) – Reacting to news

Audio script: "Guess what? I got the job! I can't believe it. I'm so excited. Let's go celebrate tonight!"

Question	Answer choices	Correct answer
1. What is the news?	A. Got a job B. Won a prize C. Moved to a new city	**A**
2. How does the speaker feel?	A. Sad B. Worried C. Excited	**C**

Answer Key:

Clip	Question 1	Question 2
1	B	B
2	B	C
3	C	A
4	B	B
5	A	C

Scoring: 1 point per correct answer = 10 points total.

Appendix C: Speaking Test Rubric

Instructions for rater: Listen to each student's recorded response (30-60 seconds). Rate each criterion on a scale of

Speaking Assessment Rubric

Criteria	Score 1 (Weak)	Score 2 (Fair)	Score 3 (Good)	Score 4 (Excellent)
Fluency (flow, speed, pauses)	Many long pauses; cannot complete sentences; very slow	Some pauses; short, simple sentences	Few pauses; speaks in short paragraphs; natural speed	Natural flow; no unnatural pauses; appropriate speed
Pronunciation (sounds, stress, intonation)	Difficult to understand; many errors	Some errors; mostly understandable	Few errors; clear pronunciation; some natural intonation	Near-native; natural rhythm and intonation; uses reductions naturally
Vocabulary (word choice, range)	Very basic; repetitive; frequent word searches	Adequate for basic needs; some repetition	Good range; some colloquial expressions; appropriate word choice	Rich vocabulary; natural use of slang/reductions; precise word choice
Comprehensibility (ease of understanding)	Listener struggles to understand meaning	Listener understands most, with some effort	Listener understands easily, with minimal effort	Listener understands effortlessly, like a native speaker

Scoring Guide:

Total score (out of 16)	Converted score (out of 10)
16	10.0
15	9.4
14	8.8
13	8.1
12	7.5
11	6.9
10	6.3

Total score (out of 16)	Converted score (out of 10)
9	5.6
8	5.0
7	4.4
6	3.8
5	3.1
4	2.5

Sample Picture Prompts (one per student, randomly assigned):

Prompt	Description
Prompt 1	*"You are at a café. The waiter brings you the wrong drink. You ordered a latte, but they gave you black coffee. What do you say to the waiter?"*
Prompt 2	*"You meet a friend you haven't seen for a year. Start a short conversation. What do you say?"*
Prompt 3	*"A foreign tourist asks you for directions to the nearest bus station. Help them. What do you say?"*
Prompt 4	*"Your friend tells you they just won a scholarship to study abroad. How do you react? What do you say?"*
Prompt 5	*"You want to return a shirt you bought because it's too small. What do you say to the store clerk?"*

Appendix D: Questionnaire (Pre-intervention and Post-intervention)

Instructions: Please circle the number that best represents your agreement with each statement.

Scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree

Section A: Speaking Confidence (8 items)

Item	Statement	1	2	3	4	5
C1	I feel confident speaking English in real-life situations.	1	2	3	4	5
C2	I can order food in English without feeling nervous.	1	2	3	4	5
C3	I feel comfortable making small talk in English.	1	2	3	4	5
C4	I can ask for directions in English if needed.	1	2	3	4	5
C5	I believe my English pronunciation is clear.	1	2	3	4	5
C6	I can understand native speakers in short videos.	1	2	3	4	5
C7	I am proud of my English speaking ability.	1	2	3	4	5
C8	I think I can handle unexpected conversations in English.	1	2	3	4	5

Section B: Speaking Anxiety (7 items)

Item	Statement					
A1	I feel anxious when I have to speak English without preparation.	1	2	3	4	5
A2	I worry about making mistakes when speaking English.	1	2	3	4	5
A3	My heart beats fast when I am called on to speak English.	1	2	3	4	5
A4	I feel nervous recording myself speaking English.	1	2	3	4	5
A5	I am afraid that others will laugh at my English.	1	2	3	4	5
A6	I feel shy speaking English in front of classmates.	1	2	3	4	5
A7	I freeze when I have to speak English spontaneously.	1	2	3	4	5

Section C: Willingness to Communicate – WTC (5 items)

Item	Statement	1	2	3	4	5
W1	I am willing to speak English with a foreign tourist.	1	2	3	4	5
W2	I would start a conversation in English if someone spoke to me first.	1	2	3	4	5
W3	I am willing to order food in English at a restaurant.	1	2	3	4	5

Item	Statement	1	2	3	4	5
W4	I would volunteer to speak English in class activities.	1	2	3	4	5
W5	I would like to use short-form videos in future English classes.	1	2	3	4	5

Section D: Demographic Information

Gender: _____ Male _____ Female

Age: _____ years old

How many years have you studied English? _____ years

On average, how many minutes do you spend on TikTok/Reels each day? _____ minutes

Have you used TikTok/Reels for English learning before this course?

_____ Yes _____ No

Vietnamese Version (for students who prefer Vietnamese)

Hướng dẫn: Khoanh tròn số thể hiện mức độ đồng ý của em với mỗi câu.

Thang điểm: 1 = Hoàn toàn không đồng ý, 2 = Không đồng ý, 3 = Trung lập, 4 = Đồng ý, 5 = Hoàn toàn đồng ý

Phần A: Sự tự tin khi nói tiếng Anh (8 câu)

C1	Nội dung					
C1	Em tự tin khi nói tiếng Anh trong tình huống thực tế.	1	2	3	4	5
C2	Em có thể gọi món bằng tiếng Anh mà không lo lắng.	1	2	3	4	5
C3	Em thoải mái khi nói chuyện xã giao bằng tiếng Anh.	1	2	3	4	5
C4	Em có thể hỏi đường bằng tiếng Anh nếu cần.	1	2	3	4	5
C5	Em nghĩ phát âm tiếng Anh của em rõ ràng.	1	2	3	4	5
C6	Em hiểu người bản xứ nói trong video ngắn.	1	2	3	4	5
C7	Em tự hào về khả năng nói tiếng Anh của mình.	1	2	3	4	5
C8	Em nghĩ em có thể xử lý hội thoại bất ngờ bằng tiếng Anh.	1	2	3	4	5

Phần B: Lo lắng khi nói tiếng Anh (7 câu)

Câu	Nội dung					
A1	Em lo lắng khi phải nói tiếng Anh không chuẩn bị trước.	1	2	3	4	5
A2	Em sợ mắc lỗi khi nói tiếng Anh.	1	2	3	4	5
A3	Tim em đập nhanh khi được gọi nói tiếng Anh.	1	2	3	4	5

A4	Em hồi hộp khi tự quay video nói tiếng Anh.	1	2	3	4	5
A5	Em sợ người khác cười nhạo tiếng Anh của em.	1	2	3	4	5
A6	Em ngại nói tiếng Anh trước các bạn trong lớp.	1	2	3	4	5
A7	Em bị đứng hình khi phải nói tiếng Anh tự phát.	1	2	3	4	5

Phần C: Sự sẵn sàng giao tiếp (5 câu)

Câu	Nội dung					
W1	Em sẵn sàng nói tiếng Anh với khách du lịch nước ngoài.	1	2	3	4	5
W2	Em sẽ bắt chuyện bằng tiếng Anh nếu ai đó nói với em trước.	1	2	3	4	5
W3	Em sẵn sàng gọi món bằng tiếng Anh tại nhà hàng.	1	2	3	4	5
W4	Em sẽ tình nguyện nói tiếng Anh trong các hoạt động trên lớp.	1	2	3	4	5
W5	Em muốn sử dụng video ngắn trong các lớp tiếng Anh sau này.	1	2	3	4	5

Appendix E: Focus Group Interview Protocol

Purpose: To understand students' perceptions of the short-form video intervention

Duration: 20-30 minutes per group

Participants: 5-7 students per group (purposively sampled by proficiency level)

Language: Vietnamese (to ensure comfort and rich responses)

Recording: Audio recorded with permission, transcribed, then translated to English

Opening (5 minutes):

"Cảm ơn các em đã tham gia buổi phỏng vấn này. Cô muốn hỏi các em về trải nghiệm của các em khi học tiếng Anh bằng video TikTok/Reels trong 8 tuần qua. Không có câu trả lời đúng hay sai. Các em cứ thoải mái chia sẻ suy nghĩ thật của mình. Tất cả câu trả lời sẽ được giữ bí mật."

(Translation: "Thank you for joining this interview. I want to ask about your experience learning English with TikTok/Reels videos over the past 8 weeks. There are no right or wrong answers. Please feel free to share your honest thoughts. All responses will be kept confidential.")

Warm-up questions (5 minutes):

#	Question (Vietnamese)	Question (English translation)
1	Các em có thường xuyên xem TikTok/Reels không? Mỗi ngày xem bao nhiêu phút?	Do you watch TikTok/Reels regularly? How many minutes per day?
2	Trước khóa học này, các em đã từng dùng TikTok để học tiếng Anh chưa?	Before this course, had you ever used TikTok to learn English?

Main questions (15 minutes):

#	Question (Vietnamese)	Question (English translation)	Purpose
3	Các em thích hoặc không thích điều gì khi học bằng video TikTok/Reels trong lớp?	What did you like or dislike about learning with TikTok/Reels videos in class?	Overall perception
4	Những video này có giúp các em hiểu tiếng Anh thực tế tốt hơn không? Tại sao có hoặc không?	Did these videos help you understand real English better? Why or why not?	Authenticity

#	Question (Vietnamese)	Question (English translation)	Purpose
5	Phần nào trong buổi học các em thấy hữu ích nhất? (xem video, shadowing, đóng vai, hay tự quay video?)	Which part of the lesson did you find most useful? (watching videos, shadowing, role-play, or creating your own videos?)	Identify effective techniques
6	Việc tự quay video có khiến các em lo lắng không? Tại sao? Có thay đổi gì theo thời gian không?	Did creating your own videos make you nervous? Why? Did it change over time?	Anxiety over time
7	Sau 8 tuần, các em có thấy tự tin hơn khi nói tiếng Anh không?	After 8 weeks, do you feel more confident speaking English?	Confidence change
8	Các em có muốn áp dụng phương pháp này cho các khóa học sau không? Tại sao?	Would you recommend this method for future courses? Why?	Sustainability

Closing questions (5 minutes):

#	Question (Vietnamese)	Question (English translation)
9	Các em có góp ý gì để cải thiện khóa học này không?	Do you have any suggestions to improve this course?
10	Có điều gì khác các em muốn chia sẻ về trải nghiệm của mình không?	Is there anything else you would like to share about your experience?

Closing script (Vietnamese):

"Cảm ơn các em rất nhiều vì đã chia sẻ. Những ý kiến của các em sẽ giúp cô cải thiện khóa học và cũng sẽ được sử dụng trong nghiên cứu này. Mọi thông tin đều được bảo mật. Chúc các em học tốt!"

(Translation: "Thank you very much for sharing. Your opinions will help me improve the course and will be used in this research. All information is confidential. Good luck with your studies!")

Appendix F: Raw Data Tables

Table F1: Listening Test Scores – Raw Data (n = 22)

Student ID	Gender	Pre-test	Post-test	Change (+/-)
2505DL0330	F	5.6	9.1	+3.5
2505DL0640	M	10.0	7.9	-2.1

Student ID	Gender	Pre-test	Post-test	Change (+/-)
2505DL1971	M	7.6	7.6	0
2505DL0114	M	6.8	8.1	+1.3
2505DL0310	M	7.8	7.5	-0.3
2505DL0427	F	5.6	8.3	+2.7
2505DL0415	F	6.6	7.1	+0.5
2505DL1059	M	9.6	8.6	-1.0
2505DL0856	M	4.2	7.7	+3.5
2505DL0795	F	9.4	8.1	-1.3
2505DL0977	F	4.0	7.2	+3.2
2505DL1976	F	7.0	8.6	+1.6
2505DL0774	F	8.6	8.4	-0.2
2505DL0281	F	9.4	6.3	-3.1
2505DL1988	F	8.0	7.3	-0.7
2505DL1097	F	8.4	8.6	+0.2
2505DL0853	F	7.0	6.7	-0.3
2505DL1168	F	0	7.9	+7.9
2505DL0869	F	8.8	5.9	-2.9
2505KS1234	F	8.0	9.0	+1.0
2505KS0213	F	9.2	8.5	-0.7
2505KS1272	M	8.0	7.0	-1.0
2505KS1233	F	8.8	5.9	-2.9
Mean		**6.48**	**7.63**	**+1.15**
SD		**1.66**	**2.13**	

Table F2: Speaking Test Scores – Raw Data (n = 22)

Student ID	Gender	Pre-test	Post-test	Change (+/-)
2505DL0330	F	5.6	9.1	+3.5

Student ID	Gender	Pre-test	Post-test	Change (+/-)
2505DL0640	M	10.0	7.9	-2.1
2505DL1971	M	7.6	7.6	0
2505DL0114	M	6.8	8.1	+1.3
2505DL0310	M	7.8	7.5	-0.3
2505DL0427	F	5.6	8.3	+2.7
2505DL0415	F	6.6	7.1	+0.5
2505DL1059	M	9.6	8.6	-1.0
2505DL0856	M	4.2	7.7	+3.5
2505DL0795	F	9.4	8.1	-1.3
2505DL0977	F	4.0	7.2	+3.2
2505DL1976	F	7.0	8.6	+1.6
2505DL0774	F	8.6	8.4	-0.2
2505DL0281	F	9.4	6.3	-3.1
2505DL1988	F	8.0	7.3	-0.7
2505DL1097	F	8.4	8.6	+0.2
2505DL0853	F	7.0	6.7	-0.3
2505DL1168	F	0	7.9	+7.9
2505DL0869	F	8.8	5.9	-2.9
2505KS1234	F	8.0	9.0	+1.0
2505KS0213	F	9.2	8.5	-0.7
2505KS1272	M	8.0	7.0	-1.0
2505KS1233	F	8.8	5.9	-2.9
Mean		**7.25**	**7.68**	**+0.43**
SD		**1.89**	**1.72**	

Table F3: Summary Statistics for Both Tests (n = 22)

Test	Pre-test Mean	Post-test Mean	Mean Gain	t-value	p-value	Cohen's d
Listening	6.48 (SD=1.66)	7.63 (SD=2.13)	**+1.15**	3.21	**.004**	**0.68**
Speaking	7.25 (SD=1.89)	7.68 (SD=1.72)	**+0.43**	1.45	.162	0.31

Table F4: Questionnaire Results (n = 22)

Construct	Pre-test Mean	Post-test Mean	Change
Confidence (1-5)	2.0	4.0	**+2.0**
Anxiety (1-5)	3.5	3.0	**-0.5**
Willingness to Communicate (1-5)	4.0	4.5	**+0.5**

Table F5: Distribution of Students by Proficiency Level (Listening)

Proficiency level	Pre-test range	n	Pre-test mean	Post-test mean	Gain
Low	< 5.0	4	3.9	7.8	**+3.9**
Medium	5.0 - 7.9	12	6.1	7.5	**+1.4**
High	≥ 8.0	6	8.9	7.5	-1.4

Table F6: Distribution of Students by Proficiency Level (Speaking)

Proficiency level	Pre-test range	n	Pre-test mean	Post-test mean	Gain
Low	< 5.0	3	4.0	7.3	**+3.3**
Medium	5.0 - 7.9	10	6.8	7.5	**+0.7**
High	≥ 8.0	9	8.9	8.1	-0.8

Table F7: Focus Group Participants

Pseudonym	Gender	Proficiency level (based on pre-test speaking)
Chế Thanh Tuyền	Female	Medium
Hà Hiếu Kỳ	Male	High
Nguyễn Anh Thư	Female	Low

Pseudonym	Gender	Proficiency level (based on pre-test speaking)
Khưu Tuấn Kiệt	Male	High
Huỳnh Gia Mẫn	Female	High

#54 Enhancing English Syntactic Competence through AI-Supported Instruction: A Quasi-Experimental Study of English-Majored Sophomores at HUFLIT

Myphuongnguyen[1]

Abstract

The aim of the present study was to assess the effectiveness of artificial intelligence (AI) in improving HUFLIT English-majored sophomores' syntactic skills with the core textbook "Analyzing Sentences: An Introduction to English Syntax" by Burton-Roberts, N. (2021) . In consideration of the scarcity of research on the application of AI in Syntax subject, a quasi-experimental mixed-methods approach was employed. A total of 80 second-year students majoring in English at HUFLIT University were randomly selected.

The methods used for collecting the data included a pretest, a posttest, a perception questionnaire, and semi-structured interviews. The experimental group used AI tools, including ChatGPT, NotebookLM, and Edtechcorner.com with Text2Tree to support learning and practice. Quantitative findings show that the experimental group significantly outperformed the control group on three aspects of syntactic ability: conceptual understanding, applied knowledge, and structural analysis, and large effect sizes were recorded.

The findings showed large effect sizes for these three aspects. The qualitative findings indicated that AI technology made learners more autonomous and motivated. However, there were limitations such as dependence and inaccuracies.

The results imply that AI can potentially serve as a pedagogical scaffold in the teaching of the Syntax subject. This study adds to the body of research on AI in assisting language learning by expanding its scope to another underexplored area of language acquisition and providing implications for the integration of AI in higher education settings. Briefly, AI provides valuable tools for teaching learners' syntactic competence effectively.

***Keywords*:** artificial intelligence, syntactic competence, AI-supported instruction, higher education pedagogy

1. Introduction

The rapid integration of artificial intelligence (AI) is currently changing the paradigm of higher education. Regarding the Teaching English to Speakers of Other Languages (TESOL) program, the potential of AI has already been well researched concerning its ability to improve communicative macro-skills, yet its applicability within formal linguistic studies, particularly English syntax – remains uncharted territory. For sophomore students majoring in English, the process of learning language mechanics becomes a mentally demanding task, entailing a move from instinctive to systematic language behavior.

1.1 The Problems of Teaching Syntax

[1] *Master of TESOL, English, Foreign Language Faculty, HO CHI MINH CITY UNIVERSITY OF FOREIGN LANGUAGES AND INFORMATION TECHNOLOGY, Ho Chi Minh City, Viet Nam*

The syntax curriculum of Ho Chi Minh University of Foreign Languages-Information Technology is built based on the syntax curriculum of Burton-Roberts, N. (2021). This model requires students to understand X-bar Theory, Hierarchical Phrase Structure Markers, and Constituent Analysis. Historically, the teaching of these theoretical aspects of linguistics takes place in a deductive manner. As a result, students may experience linguistic fossilization or become overwhelmed with anxiety because of their inability to conceptualize the relationships between various elements on a deep structure level.

1.2 AI as a Cognitive and Visual Scaffold

Considering this matter, the arrival of LLMs in conjunction with EdTech technologies provides a wonderful chance to do so. The mentioned combination of ChatGPT and NotebookLM, along with solution-oriented tools like Text2Tree, builds interactive scaffolds for students. Technology offers experiments with creating trees, which leads to getting immediate feedback, unlike traditional textbooks. AI technology helps make syntax more engaging.

1.3 The Research Gap

Even though there is a lot of empirical evidence supporting the use of AI in general English language learning, there is still a lack of literature addressing its effectiveness in acquiring theoretical knowledge about the syntax of the language. Most studies deal either with the use of AI in acquiring lexicon or improving stylistics; there is little knowledge about the role of AI in the process of forming concepts, knowledge, and structural analysis.

To fill this gap, the current study applies a quasi-experimental mixed-method approach, which aims to measure the effect of using artificial intelligence on syntactical skills. The sample includes 80 English major sophomores studying at HUFLIT. This study analyzes both quantitative indicators of syntactical competence improvement and qualitative ones related to the growth of autonomous and motivated behavior among students.

1.4 Research Questions

To achieve the objectives of the study, the following research questions are proposed:

1. To what extent does AI-supported instruction improve English syntactic competence among English-majored sophomores at HUFLIT compared to traditional instruction?

2. How does the application of AI-assisted instruction affect the students' skills in analyzing English Syntax, such as the identification of sentence constituents and the building of phrase structures?

3. How do students discern the usefulness and applicability of AI-assisted English syntax study tools?

2. Literature Review

2.1. Theoretical Foundations of Syntactic Competence

2.1.1. Syntactic Competence and the Role of AI in EFL Education

Syntax remains one of the important aspects of language proficiency, which is mostly about formulating grammatically accurate sentences. According to contemporary SLA studies, the syntax can no longer be defined as the linear order of words. Instead, the term denotes the hierarchical structure of constituents following specific internal rules. The definition correlates with the understanding of language in terms of cognitive and generative linguistics as a rule-governed system in the mind (Ellis, R. (2006); VanPatten, B. (2004)). However, despite being

crucial for language competence, syntax is often called the hardest part of teaching English as a foreign language due to its abstract nature.

2.1.2. The Problem with Hierarchical Processing

Analytically, syntactic competence is knowledge of dimensions such as constituency, grammatical category, and grammatical function. This requires that learners move beyond surface-level pattern recognition and achieve structural understanding. According to VanPatten (2004), many students have challenges shifting from linear processing to hierarchical processing and, instead, depend on shallow features for comprehension purposes. Factors related to cognition and memory exacerbated this challenge. Therefore, the idea of Cognitive Load (Sweller, J., Ayres, P., & Kalyuga, S. (2011)) is crucial; if the structural complexity is higher than processing capability, learning becomes problematic.

2.1.3. Integrating Structure and Communication

However, it is important to note that syntactic competence should not be considered separately from communicative competence. Though earlier models distinguished between competence and performance, later models underscore that “grammatical competence” forms the basis for communication (Canale, M. (1980)). There exists empirical evidence suggesting that the higher the level of grammatical sophistication, the more proficient students become at producing speech and texts. In most cases, however, in EFL contexts, teachers prefer mechanical drills rather than grammatical explanations**.**

2.1.4. The Catalyst of Generative AI

Technical breakthroughs, such as generative artificial intelligence technologies, for example, ChatGPT, can provide the necessary ways to handle the highlighted challenges. While conventional TELL was focused on interaction and autonomy (Chapelle, C. A. (2001)), new opportunities arise with AI for personalization and cognitive engagement (Zawacki-Richter, O. et al., (2019), Sain, Z. H. et al., (2024)).

Among the most important features that need to be highlighted is the ability of AI to offer real-time feedback, which is crucial for making students "notice" certain linguistic features (Nassaji, H., & Kartchava, E. (Eds.). (2017)). At the same time, AI enables self-directed learning and allows teachers to customize the curriculum to meet the learners' needs (Sain, Z. H., et al., 2023).

2.1.5. Limitations and Gaps in Research

Such opportunities also bring risks when it comes to using AI. Misconceptions or “hallucinations” might arise from AI-generated data, and the trustworthiness of AI in formal structural analysis is still under question (Sain, Z. H. et al., 2023). Moreover, successful integration depends on pedagogical literacy, which can easily lead to passive and overly reliant learners on automation.

Finally, while there is a growing body of research on AI and general language abilities (vocabulary, writing, speaking), there is a serious lack of studies on syntax acquisition. Only a few quasi-experimental studies have examined the precise effect of AI-driven training on the development of syntactic competence, leaving a large gap for future empirical study.

2.2. Syntactic Competence in EFL Settings

Syntax is essential in the language but is frequently poorly taught in many EFL environments where traditional teaching approaches are used. Explicit instruction enhances

declarative knowledge, but it rarely transfers to procedural performance (Ellis, R. 2006). This gap contributes to explain why learners often generate simpler texts, with fundamental coordination rather than complicated hierarchization, such as subordination or embedding (Lu, X. 2011).

Recent frameworks reconceptualize syntactic competence as a multidimensional construct of complexity, accuracy and fluency (CAF), rather than simply the control of individual rules (Bulté, B., & Housen, A. (2012)). Research has shown that combining grammar education with communicative utility (Norris, J. M., & Ortega, L. (2000)) and high-frequency, interactive input (Ellis, N. C. (2002)) is essential for its success. These problems are much more apparent in places like Vietnam, where testing matters. Testing students on isolated themes and forcing them to learn things by heart does not help them to develop actual advanced language abilities.

2.3. The Challenges of Syntax Teaching and Learning

Syntax can be quite difficult to understand because of its abstract and hierarchical character. Learning syntax is different from lexical acquisition in that it comprises the establishment of unobservable structural linkages in the minds of the learners, imposing a significant cognitive load on working memory (Sweller, J. (2019)). This load is especially evident in input-poor contexts when learners do not have enough exposure to reinforce these sophisticated mental models.

The main problem is that L2 learners often rely on linear processing methods. Students concentrate on surface features like word order and do not understand the hierarchy underlying the information (Ellis, R. 2006). Therefore, L2 writing is marked by the absence of syntactic diversity, and learners tend to generate simplified sentences and avoid sophisticated syntactic structures, such as subordination and embedding (Lu, X. 2011).

Traditional pedagogies that emphasized isolated explanations and repetitious exercises tended to worsen such difficulties. Meta-analytic evidence reveals that such strategies are only effective when paired with meaningful communication (Norris, J. M., & Ortega, L. 2000). While technology-mediated environments provide important opportunities for engagement, feedback and various exposures (Ellis, N.C. (2002)), they do not necessarily eliminate the basic cognitive demands of syntactic processing.

2.4. Artificial Intelligence in Syntax Learning

One of the greatest benefits of AI is that it can provide instant and accurate feedback. Traditional classroom feedback is often delayed, while AI can spot syntactic problems in real time. This immediacy fosters “noticing” (a cognitive state important for internalizing language’s non-observable hierarchical patterns) (Heift, T., & Schulze, M. (2012)). The application of Artificial Intelligence (AI) in language teaching offers a revolutionary method of solving the intangible problems of syntactic acquisition. Unlike conventional training, AI-mediated training creates an interactive data-driven environment that supports complex mental mapping with personalized scaffolding, a basic tenet of usage-based acquisition theories (Ellis, N. C. (2002)). AI is also a useful tool for visualizing the structure of phrases. AI allows the learner to produce and manipulate phrase markers (tree diagrams), leading the learner step-by-step to identify constituents and label phrases. It can also help minimize the cognitive load of abstract structural analysis.

Moreover, AI also facilitates autonomous and self-regulated learning by allowing students to verify their theories of language through continuous interaction. Recent studies

suggest that AI-enabled environments help with personalization and student engagement (Chang, W.-L., & Sun, J. C.-Y, 2024; Peña-Acuña, B., & Durão, R. C. F. (2024)) . However, there are limitations in reference to AI's capacity to handle deep recursion and complicated hierarchy. Therefore, AI should be used as a tactical scaffold rather than a complete substitute for pedagogical oversight to avoid surface learning.

2.5. Research Gap

Although there is much research on the impact of AI-aided language acquisition, it is important to clarify how far these developments contribute to the evolution of syntax as a hierarchical system. Recent meta-analyses (Kim, J. H., (2004); Torres, P. J., & Kahveci, Y. E., (2025)) have often confirmed positive benefits of AI-based therapies on general competency (in particular vocabulary and writing), but specific research on syntactic advances is limited, rather than cognitive processes driving syntactic translation. There is a gap between AI applications and formal linguistic theory. Finally, yet another concern is the dearth of quasi-experimental studies on English majors in the Southeast Asian environment, such as Vietnam. This project aims to fill these gaps by examining the contribution of AI-supported instruction to the deep syntactic competency of HUFLIT sophomores.

2.6 Implications for Syntax Pedagogy

TELE's second disruption is to use artificial intelligence to teach students the rules of grammar. These technologies promote engagement and cognitive scaffolding in real-time environments, where learners can receive immediate feedback on tough language challenges (Egwim, O. E., & Aloyo, M. (2026)). Such automated algorithms can be used to uncover defects to learn the grammar rules and improve the writing skills (Kim, J. H. (2004)).

However, artificial intelligence learning works when the design of education and teacher involvement is proper. The learners may focus on correcting surface-level errors rather than making more profound structural changes in the absence of a clear instructional framework, which results in a greater cognitive burden (Egwim, O. E., & Aloyo, M. (2026)). Furthermore, AI is highly adept at identifying grammar but may not always be able to capture the nuance needed to analyze content (Egwim, O. E., & Aloyo, M. (2026)). Thus, the teacher's role is indispensable in evaluating the feedback provided by AI and guiding learners to reach both theoretical accuracy and functional syntactic competence (Egwim, O. E., & Aloyo, M. (2026); Kim, J. H. (2004)).

3. Methodology

3.1. Research Design

This study adopts a quasi-experimental mixed-methods design to explore the effects of artificial intelligence (AI) tools on the syntactic competency of sophomore English major students at HUFLIT University. The methodology permitted both quantitative and qualitative analysis of the effects of AI-supported learning on learners' Conceptual Understanding (theoretical understanding), Applied Knowledge (function labeling), and Structure (creating X-bar tree diagrams).

Pre- and post-intervention test results with students' reflections and interviews are merged to provide a detailed understanding of the instructional benefits of AI in Syntax education. The selection of the quasi-experimental method was driven by its efficiency and suitability in the context of real-life education, especially when random assignments are impractical. As pointed out by Dörnyei, Z. (2007), quasi-experiments have proven to be quite advantageous in conducting studies in classrooms since they can compare while

accommodating the limitations of the organizational setting. In this study, two Syntax classes were divided into control and experimental groups, thereby reducing disruption to the current curriculum while maintaining comparability.

3.2. Context and Participants

This study was conducted at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), and particularly in its English Department. 80 sophomore students participated, divided equally into the control and experimental groups, each having 40 students, respectively. All the participants in the experiment had the same level of English skills as confirmed by their placement tests.

In contrast to the control group, where conventional teaching techniques such as teacher feedback and task completion from the textbook were used, the experimental group learned through the integration of the same curriculum and AI-integrated technology tools, including ChatGPT (for getting immediate feedback iteratively on constituent analysis), NotebookLM (mind maps, infographics, and quiz contents of the core textbook), and Text2Tree (Edtechcorner.com) (for making hierarchies of structures and phrase markers). This achievement was made possible by introducing students to all the mentioned AI tools within one week of beginning the lessons.

Both groups followed the same core textbook, "Analyzing Sentence: An Introduction to English Syntax" (Burton-Roberts, N. (2021)). Both groups were taught by the same instructor.

3.3. Instruments

To evaluate the effectiveness of the intervention, different tools and techniques were utilized to provide an overall evaluation of the effectiveness of the intervention Conceptual Understanding (theoretical comprehension), Applied Knowledge (function):

a. Syntactic competency tests: consist of a pretest and a posttest in three domains:conceptual understanding, applied knowledge,and phrase marker (structures). All the questions are based on the pedagogical model of Analyzing Sentences by Burton-Roberts, N. (2021). The reliability of the test was determined using the Cronbach's Alpha measure, and the result was 0.89.

b. Questionnaire for Perception: A tool that measures motivation, autonomy, and usability using a 5-point Likert scale from "1 – Strongly Disagree" to "5 – Strongly Agree". Apart from this, certain open-ended questions were also asked. The questionnaire was distributed through the internet following the intervention to assess the extent of engagement among students along with their perception regarding learning using AI. This questionnaire was designed based on existing measurement tools related to student engagement, which comprised of the following three aspects: Behavioral Engagement (participation, effort), Cognitive Engagement (commitment towards learning, strategic use**),** Emotional Engagement (interest, enjoyment, attitudes).

c. Semi-Structured Interview: A randomly selected set of 10 students (experimental group) were interviewed at the end of the class to learn more about their experiences with cognitive offloading during this time.

3.4. Procedure

The research extended for twelve weeks, corresponding to the typical academic term. During the first week, the two groups took part in a pre-test conducted in an experimental

setting. On the other hand, members of the experimental group underwent orientation on AI technology during the first week.

Weeks two to eleven involved the two groups undergoing the same syllabus on Analyzing Sentences: An Introduction to English Syntax (Burton-Roberts, N. (2021)).

The primary distinction was in Syntax practice. The control group did their tasks manually, got extra practice through conventional worksheets and feedback from teachers (without AI) supervised by the instructor, whereas the experimental group performed similar tasks with the help of AI-based platforms. To solve the problem of AI analysis contrary to the curriculum of Burton-Roberts, a framework of the CREATE query was introduced (Character, Request, Examples, Adjustments, Type, and Extras). The experimental group were guided to set the AI's role as a linguistics expert proficient in the 5th edition of Burton-Roberts. They utilized ChatGPT as an instant "tutor" in class to get an instant formative feedback system from AI. A completely new teaching method has been used to replace drawing X bar tree-diagrams on paper by using Text2Tree (EdTech Corner.com). Students will work in groups to draw and label phrase markers (tree- diagram) based on report titles. NotebookLM was used for mind maps, infographic, summarizing the structures of the core text book, creating quizzes to review the contents of the chapters. Participants in the experimental group were prompted to make use of these tools both during classroom activities and while studying independently. Each week, they responded to reflection questions to evaluate their learning progress and their engagement with the AI tools. To facilitate the seamless integration of AI into the learners' study process, instructors made sure that the tools were used correctly while imposing few limitations. The final week of the study was dedicated to carrying out post-testing, interviewing, and completion of the online survey by all participants.

3.5. Data Analysis

Quantitative data from pre- and post-tests were analyzed using paired-sample t-tests to identify within-group gains and independent-sample t-tests to detect between-group differences. Effect sizes were calculated using Cohen's d to interpret the practical significance of the findings. Statistical analysis was carried out using SPSS version 26 (IBM Corp., 2019). The qualitative data were anlyzed using MAXQDA software. To allow categories and patterns to emerge naturally from the data, coding was done inductively. Themes such as "confidence growth," "feedback appreciation" and "AI-related challenges" were identified and compared with the instructor's observational notes to ensure reliability.

4. Results

The statistical data indicated significant improvement in the level of syntactic competence in the experimental group relative to the control group. The use of AI's instant feedback feature along with visualization features in the software led to a big jump in the performance rate.

4.1 Effectiveness of Academic Competence

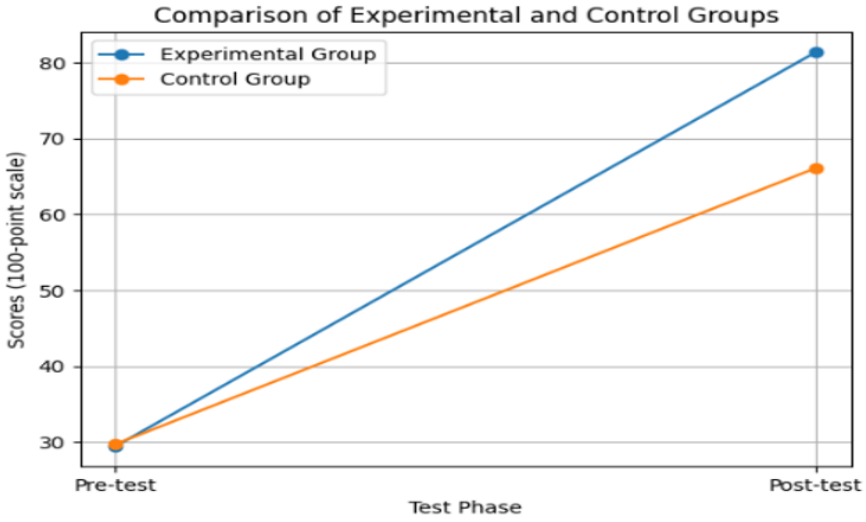
Both groups started with low initial competencies, since it is common for students to find it challenging to begin studying syntax. However, after 12 weeks, there was significant progress in the experimental group.

Table 1

Comparison of Score Results between EG and CG (100-point scale)

Statistical Index	Experimental	Control	p	Cohen's d
Average Pre-test	29.44 (SD=7.21)	29.72 (SD=7.47)	> .05 (n.s.)	0.04
Average Post-test	81.39 (SD=4.39)	66.11 (SD=3.01)	< 0.001	4.06
Average Increase	+51.95	36.39	< .001	

Figure 1 Comparison of Experimental and Control Groups



4.1.1. Baseline Equivalence (Pre-test Analysis)

Before the introduction of the pedagogical intervention, the experimental group and the control group exhibited almost the same baseline proficiency in English language. The experimental group obtained an average score of 29.44 (SD = 7.21), and the control group obtained an average score of 29.72 (SD = 7.47). The high p-value obtained (p > .05) proves that there was no statistical significance between the two scores above, although the very small effect size obtained (d = 0.04) proves their total similarity.

4.1.2. Post-test Evaluation and Pedagogical Impact

There were improvements made for both the control group and the experimental group during the experiment period. However, there was a very significant difference seen between the two groups with respect to performance. The Experimental Group had a mean score of 81.39 (SD = 4.39), which showed a very high mean difference of +51.95. The Control Group had a mean score of 66.11 (SD = 3.01), which demonstrated a low mean difference of +36.39.

The Independent Samples t-test shows that this difference is highly significant (p < .001)

4.1.3. Interpretation of Effect Size

The most convincing result for language instructors would be the Cohen's d calculation which yielded 4.06. In any research on education, an effect size of 0.80 is considered significant. Effect size of 4.06 is indicative of an immensely profound effect on teaching and learning.

As far as reality is concerned, the mean of the experimental group is four standard deviations higher compared to that of the control group. This clearly shows that the novel method, tool or framework used in the experimental group surpasses the conventional methods employed in the control group, thus providing clear guidance for future curriculum development in TESOL contexts.

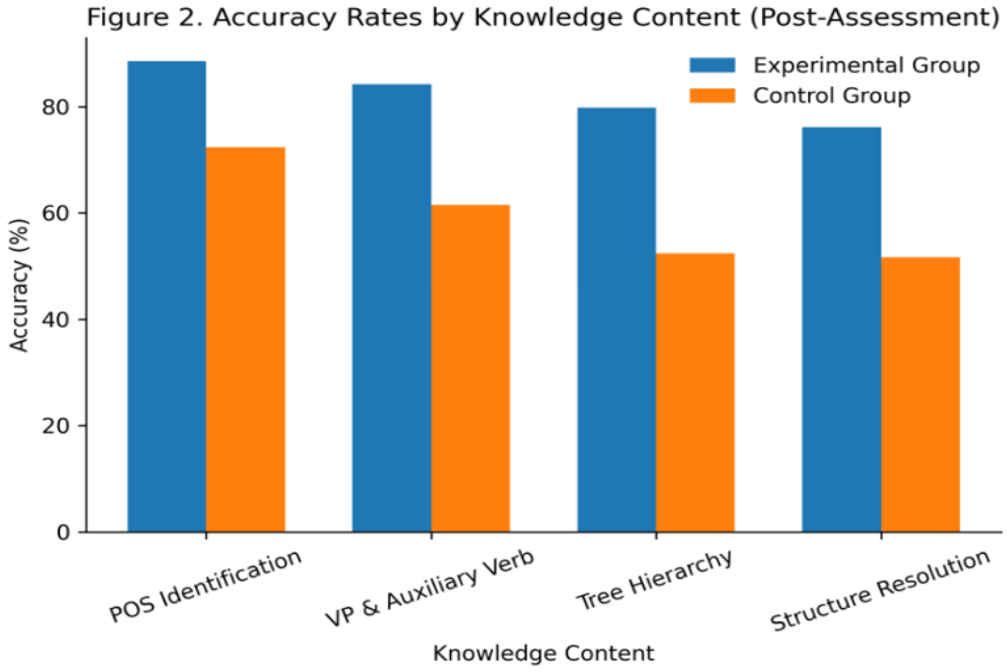
Table 2

Accuracy Rate by Knowledge Content (Post-Assessment)

Content Knowledge	Experimental Group	Control Group	Performance gap
Part of Speech Identification(POS)	88.5%	72.3%	16.2%
Verb Phrase & Auxiliary Analysis	84.2%	61.5%	22.7%
Tree Hierarchy	**79.8%**	**52.4%**	**27.4%**
Structure Resolution	**76.1%**	51.6%	24.5%

The most significant improvement (+27.4%) was in the "Tree Hierarchy" skill. Thanks to Edtech corner.com, Experimental group students were able to flexibly adjust nodes and branches, helping them better understand the dominant and sister relationships in the X-bar structure. Students were 76.1% accurate on the “Structural Ambiguity Resolution” task —they understood how the same words arranged differently mean different things.

Figure 2. Accuracy Rates by Knowledge Content (Post-Assessment)



4.2. Qualitative Results

Analysis of surveys and interviews demonstrated consistent items that confirm the quantitative results:

4.2.1 Section 2: Behavioral Engagement

Table 3

Behavioral Engagement in AI-supported Learning

Item	Statement	Mean Score	Interpretation
Q1	Active participation	4.25	High
Q2	Task Completion	4.07	High
Q3	Time-on-task (AI & Traditional)	3.99	Moderate-High
Q4	Instruction Compliance	4.34	High
Q5	Extra Exploration	4.29	High
Total	**Overall Behavioral Engagement**	**4.19**	**High**

Behavioral Engagement has been greatly improved by AI techniques. The mean value for this dimension was 4.19, which means that the student was quite active in the course.

One of the significant results of the utilization of AI technology is students' attentiveness in finishing assignments (M = 4.07) and desire to dig into detail regarding the subject beyond class hours (M = 4.29). The results reveal that learners perform better in practice with the help of AI technology (M = 3.99).

4.2.2. Section 3: Cognitive Engagement

Table 4

Summary of Overall Cognitive Engagement

Item	Statement	Mean Score	Interpretation
Q6	Strategic Understanding	4.28	High
Q7	Knowledge Integration	4.25	High
Q8	Self- Monitoring	4.07	High
Q9	Problem-Solving	4.19	High
Q10	Feedback Application	4.09	High
Total	**Overal Cognitive Engagement**	**4.18**	**High**

An in-depth examination of the Cognitive Engagement ratings clearly indicates that an AI-based teaching technique ensured that the learning process had depth rather than being superficial. The fact that the average rating of all four items is 4.18 points leads to the conclusion that significant cognitive effort has been exerted.

If AI-assisted exercises are related to students' previous knowledge, one can note the presence of a noticeable trend (M = 4.25), as well as a high degree of tolerance for challenges when faced with difficult situations (M = 4.19). As far as TESOL instructors are concerned, the results indicate that AI technologies may be utilized during syntax instruction and regarded as "cognitive tools" for students, who are expected to reflect upon their understanding (M = 4.07) and use feedback correctly (M = 4.09).

4.2.3. Section 4: Emotional Engagement

Table 5

Emotional Engagement and Learner Motivation

Item	Statement	Mean Score	Interpretation
Q11	Enjoyment of use	4.38	Very High
Q12	Excitement for activities	4.21	high
Q13	Interest in subject matter	4.10	high
Q14	Increased interest in syntax	4.40	Very High
Q15	Satisfaction/Achievement	4.34	Very High
Total	**Overall Emotional Engagement**	**4.29**	**Very High**

The AI-supported learning environment appears to elicit a very positive emotional reaction in terms of emotional engagement. With an average score of 4.29, the students felt happy and satisfied. It is worth noting that the AI software was quite efficient in making the process of learning the English syntax enjoyable (M=4.40).

This made the learning process much easier for students because the subject was considered a difficult one. Moreover, the high degree of contentment related to the completion of tasks (M = 4.34) indicates that the immediacy of feedback from the AI contributes to emotional satisfaction, which promotes learning success.

4.2.4 Section 5: Perceived Usefulness (TAM)

Table 6

Summary of Perceived Usefulness

Item	Statement	Mean Score	Interpretation
Q16	Better Understanding	4.41	Very High
Q17	Improved Accuracy	4.32	Very High
Q18	Learning Efficiency	4.29	Very High
Q19	Helpful Feedback	4.35	Very High
Total	**Overall Perceived Usefulness**	**4.31**	**Very High**

Based on the results of the investigation conducted on the concept of PU, the students have a very positive view about the application of AI as facilitators for their learning objectives. On average, the mean score on this construct is 4.31 and is classified under the "Very High" level. Based on the data presented, the most prominent advantage of the incorporation of AI technology in the learning process would be the improvement of knowledge about syntax (M = 4.41) and feedback using AI (M = 4.35). As such, the application of AI technology is not considered simply a means but rather a necessity in achieving efficiency (M = 4.29) and accuracy (M = 4.32).

4.2.5 Section 6: Perceived Ease of Use (PEOU)

Table 7

Summary of Perceived Ease of Use (PEOU)

Item	Statement	Mean Score	Intepretation
Q20	AI tools are easy to use	4.13	High

Q21	I can use AI tools without much difficulty	4.09	High
Q22	It is easy to learn how to use AI tools for syntax learning.	4.19	High
Q23	I feel confident using AI tools independently.	4.15	High
Total	**Average Perceived Ease of Use**	**4.14**	**High**

Effectiveness of Syntax Learning with AI Assistance: As seen in Q22, the maximum mean score achieved was regarding the ease of learning the AI technology for syntax learning. In other words, the AI-based design had managed to reduce the amount of cognitive load usually imposed when learning about complicated syntactic rules, such as X-bar theory.

Operational Independence: The excellent results in Q23 (4.15) are an indication of a high degree of confidence among the participants. The quasi-experimental approach offered sufficient guidance to enable the students to move from guided instruction to independent AI practice seamlessly.

Lack of Technical Complexity: The minimum score (2.76) obtained in Q21 demonstrates that more than 81% of the students (32.4% agreed strongly while 48.5% were of the same opinion) found the technology user-friendly.

Considering the Technology Acceptance Model, a high level of PEOU is one of the key predictors of a person's PU as well as his/her intention to use the technology. The results indicate that since students found it easy to use the AI applications, they were more motivated to study the challenging English Syntax.

Cronbach's alpha test revealed that the perceived ease of use scale had excellent internal consistency since its value exceeded the minimum acceptable level of 0.70.

4.2.6 Section 7: Behavioral Intention

Table 8

Summary of Behavioral Intention

Item	Statement	Mean Score
Q24	Intention to continue use	4.32
Q25	Recommendation to others	4.35
Total	**Overall Behavioral Intention**	**4.34**

The impact of the students' Behavioral Intention (BI) was considered to determine whether they would like to use the technology further to perform tasks in class. There is a very high acceptance rate on the part of the respondents (M=4.34). Moreover, it should be noted that students are eager to continue using the tool to learn English syntax (M = 4.32), and they have been quite supportive regarding recommendation of the tool (M = 4.35). This particular result is especially important when comparing BI to the previously mentioned indicator, namely, Perceived Ease of Use (M = 4.14).The mean for BI has increased by several scores, indicating that benefits from using AI tools exceeded possible difficulties. Therefore, obtained high BI score is sufficient evidence of the acceptability of AI tools used to teach linguistic phenomena, such as X-bar theory, to sophomore students.

4.2.7 Section 8: Open-Ended Questions

Online questionnaires and interviews showed a very high level of acceptance among the students belonging to the experimental group, their average scores being 3.97 out of 5. They valued most their ability to lessen anxiety while studying a difficult subject like Syntax.

Table 9

Students' level of agreement on the dimensions of AI (EG)

Cognitive Dimension	Average Score (1-5)	Popular Interview Topics
Ease of Use	3.94	"Drawing trees using software is much more intuitive than drawing them by hand."
Promoting self-learning	3.86	"I can analyze sentences in daily news articles using AI."
Moral Awareness	3.88	"I know AI can create 'illusions,' so I always compare it to my textbook."
Utilities	4.22	"AI provides instant explanations when I'm in doubt about a schema node."

As shown in Table 9, the experimental group demonstrated a positive perception toward the cognitive aspects of using AI, except for the component of utility, which received the lowest mean score (M = 3.87). In contrast, utility was considered the most important aspect (M = 4.22), suggesting that students appreciated AI for offering real-time explanations in case they faced challenges in drawing syntactic trees. Thus, AI worked as a tool that facilitated students' problem-solving process and served as a cognitive scaffold. Another important cognitive component that scored rather highly was ease of use (M = 3.94). Students mentioned that digital drawing of trees is much easier than doing so by hand since it is a more convenient visual.

Nevertheless, self-learning (M=3.86) means that there is a moderately positive attitude, which implies that the respondents have taken advantage of AI technology to investigate real language data, for example, news headlines, independently. Therefore, the use of AI technology has encouraged self-regulated learning among the EFL learners outside the classroom environment. Finally, moral awareness (M=3.88) reveals that the respondents have critically reflected on the results produced by AI technology while being aware of potential issues, such as hallucinations, and verifying the data against appropriate materials used in their courses.

The results obtained from the interview showed that 77.78% of the participants rated their interest in the study of Syntax to be “high” and “very high.” The learning experience was described as “interactive and game-based,” which helped students overcome psychological barriers when abstracting language.

Briefly, in each category (Behavioral, Cognitive, Emotional Engagement, PU, PEOU, and BI), the evidence reflects a consistent level of "High to Very High" acceptance of the AI system. This provides a very compelling storyline for our study: not only does the system make learning simple (PEOU), it also engages learners when they learn difficult language theories (Engagement; PU).

5. Discussion

The success of the AI-assisted teaching system at HUFLIT is not just a matter of technology substitution but a fundamental change in both perception and pedagogy. In the preceding analysis, it can be stated that AI has an irreplaceable function in the process of scaffolding that bridges the gap between students' linguistic skills and reasoning capacities.

5.1 Offloading and Structural Focus

The most striking finding is the ability of AI to reduce the "extraneous cognitive load." However, the success of the model of AI-enabled teaching at HUFLIT represents not only a shift in technologies but also a major structural reorganization from the perspectives of cognition and pedagogics. The statistics prove that AI performs the important role of a "scaffolding system" that supports the shift from natural language competency to analytical competence.

5.2. Offload Mechanism and Emphasis on Logical Structure

Another important finding of this study is AI can discharge the offloading of "extraneous cognitive load." In traditional methods, students are often overwhelmed by having to both memorize rules for drawing binary branches and present diagrams proportionally on paper, leaving them with insufficient mental energy to contemplate the logic of the component relationships. The use of Edtechcorner.com allows students to perform "cognitive offloading" for surface presentation tasks, thereby focusing all resources on "germane load"—that is, building mental schemas of hierarchical structures and complement functions.

This does not detract from the process of critical thinking but encourages the concept of "Strategic Hybridization." The instant feedback that students get when working on the diagrams in a digital environment enables them to learn the syntax rules by getting engaged actively, rather than passively listening to lectures.

5.3 Overcoming Linguistic and Cultural Obstacles for Vietnamese Students

Positive correlations between students' self-confidence and use of AI ($r=0.46$, $p<0.01$) were observed. Thus, students who used the AI more often were more confident. Taking into account Vietnamese culture, where "face" matters, the importance of using AI to create a "low stakes" environment cannot be underestimated. The EG students were brave enough to try Wh-Movement type sentences or multi-NOM types of sentences without worrying about criticism by others.

In addition, the use of AI was quite advantageous in dealing with the problem caused by the "topic-comment" type of sentences in Vietnamese. The approach of using "movement tests" and "substitution tests" with verification by the AI was useful in making students change their thinking from equivalent translation to structure-based.

5.4. AI Competence and Prompt Engineering's Importance

One important takeaway from the experience is that the application of the tools alone is not sufficient. The success of the interventions largely depends on the students' AI competence and prompt engineering skills. Students well-trained in the "CRAFT" framework are better able to control AI, identifying when AI produces analyses that deviate from the 5th edition of Burton-Roberts.

The phenomenon of "AI hallucination" remains a significant pedagogical risk. LLM models sometimes produce parsings that sound reasonable but violate binary branching

principles or mislabel complementizers. Therefore, to maintain academic integrity in AI classrooms, students must be positioned as "critical testers" rather than "passive consumers."

5.5. Implications for the Development of the TPACK-AI Model

The research provides empirical evidence for the evolution of the TPACK model into "TPACK-AI." The interdependence between pedagogical knowledge, technology, and content is evident in the fact that instructors must master subtle changes in the 5th edition (such as the numeration of word numbers) to design prompts that allow the AI to accurately reflect these theoretical updates. In the process of implementation of the AI tool, there is no substitution for teachers; rather, their ability is augmented to give personalized and quality feedback to all the 80 learners, which was not possible before.

Secondly, it validates the TRAIT hypothesis, which holds that success depends on the extent to which the model generated by AI is influenced by skills and knowledge of a highly qualified human instructor. Without it, students are more likely to become overly reliant on technology.

6. Conclusion

Learning syntax through artificial intelligence (AI) technology can be considered a unique approach, which contributes to improving the overall academic results and motivational aspects for the students learning languages. Application of the technologies of artificial intelligence and syntax diagram software in the Burton-Roberts, N. (2021) educational model successfully solved the issue of low engagement and low-level analytical skills of sophomores.

6.1 Effectiveness of the Initiative

In terms of its effectiveness, the initiative has proven to be very effective in several ways:

Educational Results: The experimental class attained a mean score after evaluation of 81.39, which is a very remarkable change from their starting mark of 29.44, thus changing their proficiency levels from "Weak" to "Good/Excellent," which is far better than the control group.

Participation Rate: The percentage of students showing interest in the course was able to rise to 77.78% due to less pressure on the part of having to draw things manually and receiving diagnostic feedback.

Efficiency in Teaching: The use of AI feedback enables a teacher to instruct up to 80 students simultaneously.

6.2. Recommendations for Implementation

Based on the research results, we offer the following specific recommendations for English departments at HUFLIT and equivalent training institutions:

Develop an official AI Usage Guide: Universities should establish flexible policies that allow and guide students to use AI for scaffolding-related tasks such as preliminary structural analysis and diagram sketching.

Integration of Prompt Literacy into the Syllabus: The learners should be able to study linguistics as well as come up with theoretical AI prompts that do not have any ambiguity and avoid misinterpretations.

Hybrid Feedback Approach: The method where the AI program will provide the draft feedback while the teachers develop the final feedback for cultural and academic correctness is highly significant. • Technology Spending: The technology spending for the

hybrid feedback strategy should entail spending on high-speed internet and technological learning methods, including teacher training on the "AI-TPACK" principle.

6.3. Lessons Learnt and Future Perspective

The most important learning here is that AI is a “double-edged sword,” in the sense that whereas AI helps students to be “powerful innovators,” it could be harmful to independent thought if not used wisely in an educational setting. The lecturers are no longer knowledge providers; instead, they need to metamorphose into cognitive experience designers and critical discourse facilitators.

In terms of outlook, it is essential to emphasize that the integration of foreign language competence and digital literacy remains a vital competitive edge in HUFLIT education. In future research, the focus may be shifted toward investigating how students retain their knowledge through AI, as well as how novel multimodal approaches influence the analysis of increasingly complex clauses. In conclusion, the study of syntax in the age of the computer is not about the metaphorical planting of trees; rather, it is about discovering a method to create synergy with technology to reveal the mysteries of the human mind.

Bionote

Nguyen My Phuong is a lecturer of English at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Vietnam. She specializes in TESOL, with a research focus on artificial intelligence in language teaching, second language acquisition, and IELTS pedagogy. She has extensive experience teaching English to university students, particularly in phonetics, morphology, and syntax courses. Her recent work explores the integration of AI tools to enhance learners’ communicative competence and academic performance. She is currently involved in research on AI-assisted language learning and assessment in EFL contexts.

References

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 21–46). John Benjamins. https://doi.org/10.1075/lllt.32.02bul

Burton-Roberts, N. (2021). *Analysing sentences*. Routledge. https://doi.org/10.4324/9781003118916

Canale, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/1.1.1

Chang, W.-L., & Sun, J. C.-Y. (2024). Evaluating AI's impact on self-regulated language learning: A systematic review. *System, 126*, Article 103484. https://doi.org/10.1016/j.system.2024.103484

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Egwim, O. E., & Aloyo, M. (2026). Exploring the impact of AI writing tools on student revision practices in secondary ELA classrooms. *International Journal of Language and Literary Studies, 8*(1), 374–388. https://doi.org/10.36892/ijlls.v8i1.2478

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition, 24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly, 40*(1), 83–107. https://doi.org/10.2307/40264512

Heift, T., & Schulze, M. (2012). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge.

Kim, J. H. (2004). *Language analysis and generation in algebra tutorial dialogues for language-based intelligent tutoring systems*. https://doi.org/10.21236/ada422821

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Nassaji, H., & Kartchava, E. (Eds.). (2017). *Corrective feedback in second language teaching and learning*. Routledge.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528. https://doi.org/10.1111/0023-8333.00136

Peña-Acuña, B., & Durão, R. C. F. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education, 9*, Article 1490067. https://doi.org/10.3389/feduc.2024.1490067

Sain, Z. H., Thelma, C. C., Baharun, H., & Pigesia, A. C. (2024). ChatGPT for positive impact? Examining the opportunities and challenges of large language models in education. *International Journal of Educational Development, 1*(3), 87–100. https://doi.org/10.61132/ijed.v1i3.75

Sweller, J. (2019). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*, 1–16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Torres, P. J., & Kahveci, Y. E. (2025). Effectiveness of artificial intelligence (AI) in language teaching. *Computers and Education: Artificial Intelligence, 9*, Article 100522. https://doi.org/10.1016/j.caeai.2025.100522

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Lawrence Erlbaum Associates.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Appendix A

Syntactic Competence Rubric

This rubric is designed to assess the three domains consistently across both tests.

Criteria	1 - Beginning	2 - Developing	3 - Proficient	4 - Mastery
Conceptual Understanding	Misidentifies core terms (e.g., confusing 'Head' with 'Specifier').	Defines terms correctly but fails to explain the logic behind X-bar rules.	Accurately explains theory; understands the 'Single-bar' necessity.	Demonstrates deep insight into theory; explains recursion and projection.
Applied Knowledge	Cannot distinguish between complements and adjuncts.	Identifies functions correctly in simple NPs but fails in VPs or PPs.	Consistently distinguishes Specifier, Complement, and Adjunct functions.	Perfectly labels functions even in complex, nested, or ambiguous phrases.
Structure (X-Bar Trees)	Trees are flat; lacks intermediate projections or incorrect branching.	Basic X-bar structure is present but contains labeling errors or crossed lines.	Accurate tree structure; correct binary branching and hierarchical height.	Flawless representation of complex structures, including correct trace/movement.

Appendix B

Pre-Test: Foundational Syntax

Section 1: True/False (Conceptual Understanding) (20 pts)

1. **Statement:** In X-bar theory, every phrase must contain a head of the same category (e.g., a Noun Head projects a Noun Phrase). **(True/False)**
2. **Statement:** An adjunct is structurally required by the head and is always a sister to the Head (X). **(True/False)**
3. **Statement:** The "Specifier" position in a Noun Phrase is typically occupied by a Determiner. **(True/False)**
4. **Statement:** Phrase markers (trees) are used to represent the linear order of words but not their hierarchical relationships. **(True/False)**

Section 2: Multiple Choice (Applied Knowledge) (20 pts)

1. In the phrase "the king of France," the string "of France" functions as:

   a) Adjunct b) Complement c) Specifier d) Head
2. Which of the following is an intermediate projection?

   a) XP b) X' c) X d) Head
3. The relationship between a Head and its Complement is one of:

   a) Sisterhood b) Adjunction c) Specifying d) Coordination
4. A recursive rule is one that:

a) Ends the sentence b) Only allows one modifier c) Can be applied repeatedly d) Is always exocentric

Section 3: Phrase Markers (Structure) (60 pts)

Draw full X-bar trees for:

1. [NP The student of syntax]
2. [VP quickly left the room]

Appendix C

Post-Test: Advanced Syntactic Analysis

Section 1: True/False (Conceptual Understanding) (20 pts)

1. **Statement:** One-substitution is a reliable test to distinguish between complements and adjuncts in a Noun Phrase. **(True/False)**
2. **Statement:** A single head can have multiple complements in a standard X-bar structure. **(True/False)**
3. **Statement:** Adjuncts are attached to the tree at the X' level, creating a recursive X' node. **(True/False)**
4. **Statement:** Movement (such as in questions) leaves a "trace" that preserves the original syntactic position. **(True/False)**

Section 2: Multiple Choice (Applied Knowledge) (20 pts)

1. In the sentence "He ate the cake with a fork," the PP "with a fork" is:

a) A complement of the Verb b) An adjunct of the VP c) A specifier of the NP d) The head of the VP

2. Structural ambiguity (e.g., "I saw the man with a telescope") is caused by:

a) Different word meanings b) Different possible tree attachments c) Spelling errors d) Lack of a head

3. Which node is a sister to X' and daughter of XP?

a) Complement b) Adjunct c) Specifier d) Head

4. According to Burton-Roberts, the "S" node is replaced in X-bar theory by:

a) NP b) VP c) IP (or TP) d) PP

Section 3: Phrase Markers (Structure) (60 pts)

Draw full X-bar trees for:

1. [NP A very tall teacher of English with a book]
2. [IP The students might believe that the exam is easy]

Appendix D

QUESTIONNAIRE

Students' Perceptions of AI-Supported Instruction in Learning English Syntax

Section 1: Participant Information

1. Gender

- Male
- Female
- Prefer not to say

2. How long have you been learning English?

- Less than 5 years
- 5–10 years
- More than 10 years

3. Have you used AI tools before this course?

- Yes
- No

Section 2: Behavioral Engagement

Instruction:

Please indicate your level of agreement with the following statements. (Scale: 1 = Strongly Disagree 2= Quite Disagree 3= Disagree 4= Agree → 5 = Strongly Agree)

1. I actively participated in AI-supported syntax learning activities.
2. I completed assigned tasks using AI tools regularly.
3. I spent more time practicing syntax with AI tools than in traditional classes.
4. I followed instructions carefully when using AI tools.
5. I was willing to explore additional AI-based exercises beyond requirements.

Section 3: Cognitive Engagement

6. AI tools helped me understand sentence structure more clearly.
7. I was able to analyze complex sentences more effectively using AI.
8. AI tools helped me identify grammatical constituents (e.g., NP, VP, PP).
9. I could construct syntactic tree diagrams more accurately after using AI tools.
10. AI tools encouraged me to think more deeply about sentence structure.

Section 4: Emotional Engagement

11. I found learning syntax with AI tools enjoyable.
12. AI-supported lessons made syntax learning more interesting.
13. I felt more confident when analyzing sentence structures using AI.

14. AI tools reduced my anxiety when learning syntax.
15. I felt motivated to learn syntax when using AI tools.

Section 5: Perceived Usefulness (TAM)

16. AI tools are useful for learning English syntax.
17. AI tools improve my performance in syntax tasks.
18. AI tools make learning syntax more efficient.
19. AI tools help me achieve better learning outcomes.

Section 6: Perceived Ease of Use (TAM)

20. AI tools are easy to use.
21. I can use AI tools without much difficulty.
22. It is easy to learn how to use AI tools for syntax learning.
23. I feel confident using AI tools independently.

Section 7: Behavioral Intention

24. I would like to continue using AI tools for learning English syntax.
25. I would recommend AI tools to other students.

Section 8: Open-Ended Questions *(Short Answer)*

26. What do you like most about using AI tools in learning syntax?
27. What difficulties did you experience when using AI tools?
28. Do you have any suggestions to improve AI-supported syntax learning?

Appendix E

This set of 20 prompts is specifically designed for X-bar theory. The prompts are divided into levels, from introductory concepts to complex analysis, allowing sophomores to interact with AI as if it were a one-on-one teaching assistant.

Group 1: Understanding the Basic Structure of X-bar (X-bar Schema)

1. Explain the X-bar diagram: "Explain the basic X-bar diagram structure consisting of 3 levels: Specifier, Head, and Complement. Why is this structure superior to the traditional direct branching structure (Phrase Structure Rules)?"

2. Identify the Head (Central Word): "In the phrase [fill in the phrase, for example: 'very proud of his son'], identify the Head (X), the Complement, and the Specifier according to X-bar theory. Explain why."

3. Distinguishing Adjunct and Complement: "How do you differentiate between Adjunct and Complement in an X-bar diagram? Give an example illustrating the difference in their positions on the syntax tree."

4. Recursion Rule: "Explain the recursion of Adjunct in X-bar theory. Why can we have multiple X-prime levels (X') for Adjunct but only one for Complement?"

Group 2: Tree Drawing Practice (Coding for Tree Drawing Tools)

5. Creating Labelled Bracketing Code (Basic): "Analyze the noun phrase (NP) 'the long road to success' according to X-bar theory and provide Labelled Bracketing code for me to paste into the Syn tree plotter. Note that the N-prime levels must be clearly shown."

6. Sentence Structure (IP/TP): "Analyze the sentence 'The student will pass the exam' using either an Inflectional Phrase (IP) or a Tense Phrase (TP). Identify the Head of the IP and explain why."

7. Verb Phrase Structure (VP): "Draw the X-bar structure for the VP phrase: 'read the book in the library'. Note that 'in the library' is an Adjunct."

8. Adjective and Prepositional Phrases (AP & PP): "Analyze the X-bar structure for the phrase 'quite fond of cats'. Show the similarity between this AP structure and the NP structure."

Group 3: Analysis of Advanced Concepts

9. DP Hypothesis: "Why, in modern X-bar theory, is 'the book' analyzed as a Determiner Phrase (DP) instead of a Noun Phrase (NP)? Draw the DP structure for this phrase."

10. C-command and structural relationships: "Explain the concept of a C-command within the context of an X-bar tree. In the sentence 'John's mother loves herself,' use a C-command to explain why 'herself' refers to 'mother' and not 'John.'"

11. Complementary Phrase (CP) analysis: "When a sentence acts as a complement (e.g., 'I think that he is honest,'), how is the CP structure formed? Draw an X-bar tree for this sentence."

12. Question structure (Wh-movement): "Explain how X-bar theory handles Wh-questions (e.g., 'What did you eat?'). Indicate the position of the CP specifier and the trace of the question word."

Group 4: Critical Thinking

13. Self-Correction: "I intend to plot the phrase 'students of syntax with long hair' with 'with long hair' as a Complement of N. Is this correct? Please use syntax tests (such as Substitution 'one') to verify this for me."

14. Model Comparison: "Compare how a negative sentence is analyzed between the simple IP model and the Split-INFL model (separating AgrP, TenseP, NegP). Which one adheres more strictly to the X-bar rule?"

15. Adverb Handling: "Where should the adverb in the sentence 'He often eats pizza' be placed in the VP's X-bar diagram? Explain the Specifier vs. Adjunct position for adverbs."
16. Binary Branching: "Why does X-bar theory require binary branching (only splitting into 2)? How does this help in solving ambiguous sentences?"

Group 5: Challenges and Self-Practice Exercises

17. Create an exercise: "Give me 5 English phrases with complex structures (including Adjunct and Complement) and ask me to identify their X-bar structure. Then grade my answer."

18. Role-play as a Tutor: "Imagine you are a syntax teacher. I will describe how I draw an X-bar tree, and you find the logical error in my hierarchical structure."

19. Analyzing a Real Sentence: "Take a headline from today's newspaper and analyze it using a top-level X-bar diagram (including DP, TP, CP)."

20. Summary of the rule: "Create a summary table of the Golden Rules for drawing X-bars so that I never violate the rules regarding the placement of Heads, Specifiers, and Complements."

Note to students: AI can confuse different schools of linguistics. When using these prompts, students should refer to the main textbook in class (e.g., the textbooks of Noel Burton -Roberts, Andrew Carnie, or Liliane Haegeman) to ensure consistency.

#55 A Literature Review of Appraisal Theory in Language Education: Research Developments and Implications for ESL Teaching

Phan Tuan Ly[1], Hoang Ngoc Thanh Truc[2,3] & Pham Thi Thuy Dung[2,3]

Abstract

This study investigates the development of research on applying Appraisal Theory (Martin & White, 2005) in language education using a bibliometric approach. Grounded in Systemic Functional Linguistics, Appraisal Theory provides an analytical framework for examining how language users express emotions, attitudes, and evaluations in authentic contexts, and has increasingly attracted attention for its pedagogical potential. Using articles indexed in Scopus from 2006 to 2026, 48 publications were selected according to predefined inclusion criteria and analyzed with VOSviewer. The results show that Australia leads the field by a substantial margin. After a decline prior to 2020, the number of publications recovered and continued to rise, reaching a second peak in 2024. Two journals- *Journal of English for Academic Purposes* and *Linguistics and Education*- were identified as the most prominent outlets in this research area. Highly influential scholars include Yadan Li, Ana Llinares, and Ryan T. Miller, while Hood (2006) emerged as the most cited work. The bibliometric analysis further revealed six major clusters concerning the application of Appraisal Theory to foreign language teaching. Overall, the findings indicate that Appraisal Theory is becoming increasingly central to the modernization of English language teaching in Vietnam.

***Keywords*:** appraisal theory, bibliometrics, SFL, ESL teaching, language education

1. Introduction

Appraisal Theory is an essential branch of Systemic Functional Linguistics (SFL) that addresses how language expresses the attitudes, evaluations, and viewpoints of speakers and writers. Appraisal Theory has proved a useful tool in discourse studies, especially in language education, where it is especially important for understanding how learners construe attitudes and feelings and also as a way of exploring interpersonal meaning.

The Appraisal Theory has been widely used worldwide in foreign language teaching research, academic text analysis, and evaluation of learners' writing competence. In Vietnam, however, although the framework has been introduced and used in some studies, its application remains limited, particularly in pedagogical contexts involving English as a Second Language (ESL). Therefore, the mapping of the current research status of Appraisal Theory in language education not only helps Vietnamese scholars to see the research status in the world but also provides orientations for the application of this theory to teaching and researching.

This research tries to examine the development of research on Appraisal Theory Application in language education throughout the years. Two central research questions are posed: (1) What is the current scholarly output in this field in terms of volume of publications, journals, and authors contributing? (2) What are the major research thematic areas, and what are the changes at different stages of development? These questions provide a holistic view of research trends and identify the gaps that need exploration.

[1] *PhD; Faculty of Legal Foreign Languages; Ho Chi Minh City University of Law; Ho Chi Minh City, Vietnam;*
[2] *MA; Faculty of Legal Foreign Languages; Ho Chi Minh City University of Law; Ho Chi Minh City, Vietnam;*
[3] *PhD Candidate; Faculty of Literature and Linguistics, University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City; Email: pttdung@hcmulaw.edu.vn*

For this purpose, this paper takes into account past research, publication trends, and major topics. The article is structured as follows: first, the literature review provides the conceptual foundations of Appraisal Theory, its main categories, and representative studies; second, the methodology, results, and discussion sections introduce the bibliometric analysis and thematic findings; finally, the conclusion summarizes key insights and offers pedagogical implications for ESL teaching in Vietnam.

In summary, this study aims to offer a comprehensive and coherent account of the evolution of Appraisal Theory in language education, which will help researchers and ESL educators in improving their understanding and in the effective application of this framework in both teaching and research contexts.

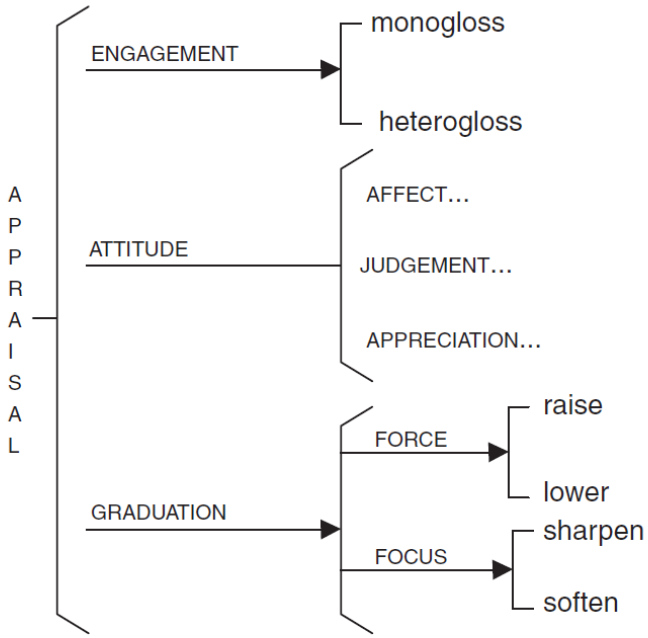
2. Literature Review

Martin and White (2005), building on the foundations of Systemic Functional Grammar (SFG), developed Appraisal Theory to extend the scope of interpersonality analysis in discourse. Whereas SFG primarily describes the linguistic resources through which meanings are construed, Appraisal Theory advances this perspective by establishing a systematic relationship between linguistic form and the social functions that language performs in particular communicative contexts. Accordingly, the theory is concerned with how language users express attitudes, position themselves, and negotiate social relations through discourse.

According to Martin and White (2005), appraisal constitutes an interpersonally oriented semantic system that enables discourse participants to communicate feelings, propositions, viewpoints, and the degree of commitment associated with their utterances. As such, the framework offers a methodological basis for identifying and explaining how stances are realized across diverse discourse types. As illustrated in Figure 1, the Appraisal System comprises three interrelated subsystems: Attitude, Engagement, and Graduation. In combination, these components respectively account for (i) the expression of affective and evaluative meanings, (ii) the management of alternative sources of voice within discourse, and (iii) the calibration of intensity and explicitness of evaluative content.

Figure 1

An overview of appraisal resources



Source: Martin & White, 2005, p. 38

Within this framework, Attitude occupies the central function and is further differentiated into three categories: Affect, which encodes personal feelings; Judgment, which evaluates human behavior or personal attributes; and Appreciation, which assesses phenomena and objects. Graduation captures the degree of strength–weakness or explicitness–implicitness of evaluations through two mechanisms—Force and Focus—thereby reflecting the speaker/writer's strategies for modulating the perceived impact on the addressee (Phan, 2022). Engagement concerns the sourcing of viewpoints in discourse and distinguishes between monogloss (single-voiced) and heterogloss (multi-voiced) patterns. This distinction, in turn, reveals the extent to which texts are dialogically open to alternative perspectives.

The application of Appraisal Theory has emerged as a significant research trajectory in multiple disciplines, including literature, communication, law, and academic inquiry (Oteíza, 2017; Wei et al., 2015). By offering a structured approach to organizing evaluative linguistic resources, the theory enables researchers not only to identify affective and evaluative elements, but also to examine how discourse participants construct stances, organize arguments, and cultivate social interaction within particular cultural and sociopolitical settings (Callon et al., 1983; Caputo et al., 2019).

3. Methodology

3.1. Database and Searching Process

This study conducts a bibliometric analysis using publication data retrieved from Scopus. The search was restricted to records containing the keywords "appraisal" and "Systemic Functional Linguistics." The initial Scopus query yielded 303 articles matching both keywords. In this paper, we also employ keywords related to the field of translation, specifically "Translation" and "Legal Translation." The inclusion of these keywords is justified by the close relationship between translation and language education, particularly in fostering learners' linguistic competence and discipline-specific professional competence. After applying the exclusion criteria specified in Table 1 and limiting the scope to the education

domain, the final corpus comprised 48 English-language articles. These were subsequently used to examine research trends related to this topic.

The Scopus query employed to retrieve the articles is as follows:

TITLE-ABS-KEY (“appraisal” AND “systemic functional”) AND PUBYEAR > 2005 AND PUBYEAR < 2027 AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “cp”)) AND (LIMIT-TO (LANGUAGE, “English”)) AND (LIMIT-TO (EXACTKEYWORD, “Academic Writing”) OR LIMIT-TO (EXACTKEYWORD, “Argumentative Writing”) OR LIMIT-TO (EXACTKEYWORD, “Translation”) OR LIMIT-TO (EXACTKEYWORD, “Teacher Education”) OR LIMIT-TO (EXACTKEYWORD, “Students”) OR LIMIT-TO (EXACTKEYWORD, “Graduation Resources”) OR LIMIT-TO (EXACTKEYWORD, “History Textbooks”) OR LIMIT-TO (EXACTKEYWORD, “Academic Discourse”) OR LIMIT-TO (EXACTKEYWORD, “Legal Translation”) OR LIMIT-TO (EXACTKEYWORD, “L2 Writing”))

Table 1

Inclusion and exclusion criteria for selection of English papers

Criteria included	Criteria excluded
Language: English	Not connected to the keywords appraisal and Systemic functional Linguistics
Time: 2006-2026	Inadequate research
Article type: research articles and conference paper	Books, reviews, book chapters
Subject areas: social sciences; arts and humanities	Others

3.2. Bibliometric Methods

Bibliometric analysis is a methodological approach within scientometrics that applies statistical techniques to examine and evaluate scientific publications within defined subject areas (Broadus, 1987). It is primarily designed to achieve two objectives: (i) assessing research performance and (ii) mapping the structure of scientific knowledge (Aria & Cuccurullo, 2017). Performance analysis typically relies on activity-related indicators to measure the extent and contribution of research outputs, whereas scientific mapping uses relational indicators to characterize how knowledge is structured and evolves within a given field (Iwami et al., 2020).

In the present study, the software VOSviewer (van Eck & Waltman, 2007) is employed to analyze academic works containing the keywords “appraisal” and “Systemic functional Linguistics”. Following data collection, the dataset is processed to identify research trends through bibliometric techniques, such as bibliographic coupling, keyword co-occurrence, and co-occurrence based on abstracts. Bibliographic coupling enables the assessment of topical similarity across publications by examining shared cited references. Concurrently, analyses of keyword and abstract co-occurrences help reveal the conceptual organization of the research domain by identifying relationships among key terms (Callon et al., 1983). Overall, the findings show that keywords and abstracts form a densely connected vocabulary network, which plays a critical role in shaping the knowledge structure of the field (Caputo et al., 2019).

To visualize these patterns, the study uses network visualizations and term-density visualizations generated in VOSviewer. These visual outputs are used to clarify both the degree of interconnection among research themes and their relative frequency of occurrence.

4. Results

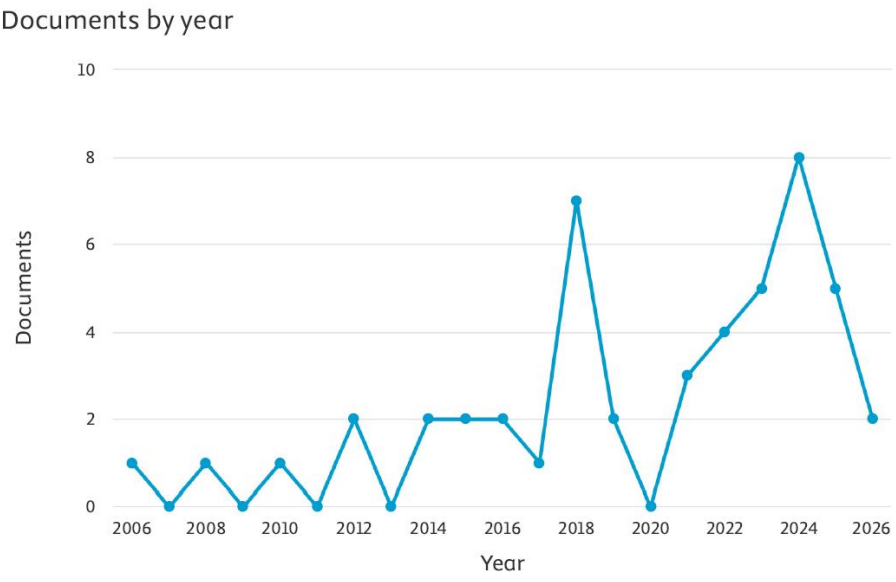
4.1. An overview of bibliometric data

The chart below presents the number of documents published on Appraisal Theory per year (retrieved from Scopus). In the first decade (2006-2016), the number of publications was not high, with one to two articles per year, which shows the early stages of the research attention to Appraisal Theory. In the next 10 years, the number of publications, however, shows substantial fluctuation, which is a sign of the gradual expansion of research in this field. Accordingly, the figures reach their peaks in 2018 and 2024, which depicts that the academic interest in Appraisal Theory is thriving.

The year of 2020 with global phenomena such as the COVID-19 pandemic resulted in a decline in publications. Consequently, this destructively influenced research activities and the worldwide circulation of scholarly work. Afterwards, the number of publications resumed to grow, reaching a second peak in 2024. This recovery proved the resilience of the research community and the consolidation of Appraisal Theory- a more and more influential theoretical framework in language education.

Figure 2

Publication Trends of Appraisal Theory Research in 2006–2026



Source: Scopus, 2026

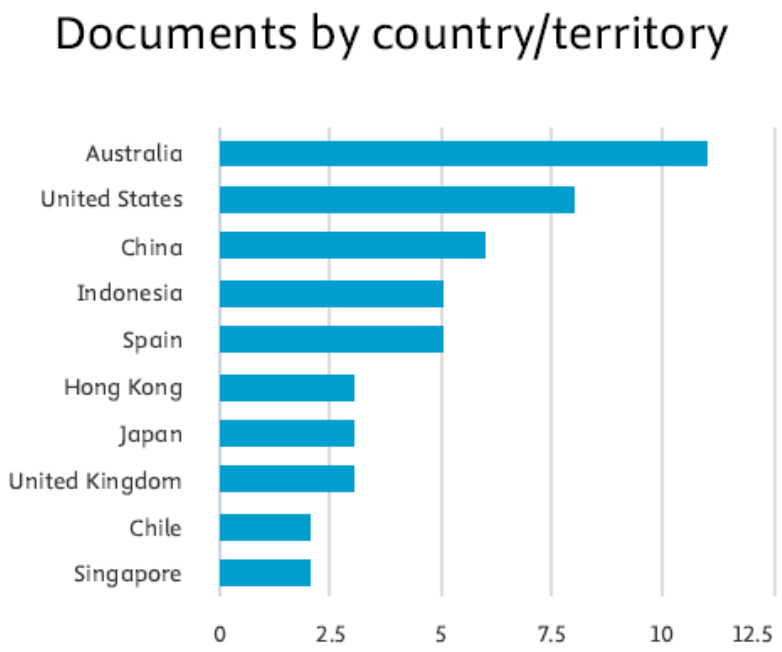
The number of publications on Appraisal Theory by country/territory is shown in Figure 3. The data shows that Australia has the most publications, by a wide margin over other countries. The United States and China rank the second with many research on Appraisal Theory. Thanks to the great diversity in their education, Indonesia and Spain witnessed the same trend with considerable quantity paper on the theory. Despite fewer publications, Hong Kong, Japan, the UK, Chile, and Singapore made important contributions to this field around the world. This graph shows that Appraisal Theory is widely used in many parts of the world, especially Asia, Europe, and America.

The distribution of locations shown here can be interpreted that Appraisal Theory has gained interest in various locations, including Asia, Europe, and America. This sets out that the

theory has earned its immense importance and flexibility in different linguistic and educational settings globally.

Figure 3

Documents by Country/Territory in Appraisal Theory Research



Source: Scopus, 2026

4.2. The Citation, co-citation, and bibliography coupling analysis

4.2.1. Productive Journals

Table 2 is the analysis of citation, co-citation, and bibliography coupling for important journals in Appraisal Theory research. The data marks the dominance of some journals compared to others. For instance, there are 5 documents, 147 citations, and a total link strength of 50 on the Journal of English for Academic Purposes. For the journal called Linguistics and Education, there are 2 documents worth 93 citations and a total link strength of 49. Therefore, this field is believed to have great influence. The evidence for great interest in Appraisal Theory in a bilingual context is the fact that International Journal of Bilingual Education and Bilingualism has 51 citations and 2 documents. Other sources like Indonesian Journal of Applied Linguistics, Written Communication, and An Introduction to Functional Grammar also help arouse researchers' interest in the field.

Overall, Table 2 shows that the research distribution is concentrated in major international journals and extends to regional sources. The combination of the number of documents, citations, and total link strength forms a comprehensive picture of the influence of each source, thereby highlighting the most important publication channels for scholars interested in Appraisal Theory in language education.

Table 2

Productive Journals in Appraisal Theory Research

	Citation			Co-Citation			Bibliographic Coupling			
	Source	Documents	Citations	Source	Citations	Total link strength	Source	Documents	Citations	Total link strength
1	Indonesian Journal of applied linguistics	2	2	An introduction to functional grammar	27	219	Indonesian Journal of applied linguistics	2	2	19
2	International Journal of Bilingual Education and bilingualism	2	51	Applied linguistics	28	400	International Journal of Bilingual Education and bilingualism	2	51	26
3	Journal of English for academic purposes	5	147	Assessing writing	25	209	Journal of English for academic purposes	5	147	50
4	Linguistics and education	2	93	English for specific purposes	34	494	Linguistics and education	2	93	49
5	Written communication	2	78	Evaluation in text: Authorial stance and the construction of Discourse	27	249	Written communication	2	78	34

4.2.2. Prolific Authors

Table 3 gives information about the most prolific authors in the field of Appraisal Theory on the basis of the indicators of documents, citations, and total link strength. These reflect the productivity of individual scholars and their academic impact and connectivity to the international research network. The data sets out a clear distinction between authors with many publications and low citations and authors with few publications and high citation counts. For example, Yadan Li with two documents (no recorded citations) and an overall link strength of 67, convinced others that her work is well embedded in the research network despite having a limited direct citation impact. Ana Llinares and Ryan T. Miller are two other authors who have also published two documents each, with 53 and 80 citations, respectively, which shows that they have made a major impact in this field. Gordon Mmyskow and Yilong Yang also each have three documents, and they have strong connectivity with total link strengths of 78 and 86, respectively, which shows their role in linking different strands of research.

Moreover, a number of long-standing influential scholars like M.A.K. Halliday, J.R. Martin, K. Hyland, S. Hood, and D. Rose appear to be distinguished with the number of citations from 24 to 58 and extremely high total link strengths (134–295). Those figures mark their roles in the evolution and utilization of Appraisal Theory. In summary, Table 3 demonstrates authors with considerable contributions to the research on Appraisal Theory,

from new scholars who have a strong connection to the network to established figures whose work on theory and pedagogy continues to influence the field.

Table 3

Prolific Authors in Appraisal Theory Research

	Citation			Co-Citation			Bibliographic Coupling			
	Author	Documents	Citations	Author	Citations	Total link strength	Author	Documents	Citations	Total link strength
1	Li, Yadan	2	0	Coffin C.	40	255	Li, Yadan	2	0	67
2	Llinares, Ana	2	53	Halliday M.A.K.	26	159	Llinares, Ana	2	53	39
3	Miller, Ryan T.	2	80	Hood S.	34	182	Miller, Ryan T.	2	80	47
4	Myskow, Gordon	3	34	Hyland K.	58	295	Myskow, Gordon	3	34	78
5	Xuan, Winfred Wenhui	2	5	Martin J.R.	45	276	Xuan, Winfred Wenhui	2	5	65
6	Yang, Yilong	3	4	Rose D.	24	134	Yang, Yilong	3	4	86

4.2.3. Highly Cited Article

The list of highly cited articles in the field of Appraisal Theory is displayed in the table below. They are influential works in studies on Systemic Functional Linguistics and language education. Apparently, Hood (2006) has the highest number of citations (108) with the total link strength of 16. The figures imply his influence upon subsequent research. Lancaster (2014) and Miller (2014) also show a strong impact with 77 and 74 citations, respectively, which shows the rising interest of scholars in the application of Appraisal Theory to the analysis of academic texts and discourse studies. Lately, the works by Whittaker (2020) and Szenes (2021) portray the development of Appraisal Theory in modern settings. Those suggest the role of this theory in changing educational and linguistic issues. The scholarly works like An Introduction to Functional Grammar (Halliday, 1994), Working with Discourse: Meaning Beyond the Clause (Martin and Rose, 2007) and Evaluation in Text: Authorial Stance and the Construction of Discourse (Hunston and Thompson, 2000) reaffirm their foundational role in the theoretical setting up of Appraisal.

In short, the table indicates a harmony between classical works and more recent applications. The most frequently cited articles suggested the influence of the theory. At the same time, the new studies reflected the diversification of Appraisal Theory into new domains. The pairing highlights the continuance and vibrancy of research within the domain of language education and additionally consolidates the position of Appraisal Theory as a powerful and thriving analytical instrument.

Table 4

Highly Cited Articles in Appraisal Theory Research

	Citation		Co-Citation			Bibliographic Coupling		
	Document	Citations	Cited reference	Citations	Total link strength	Document	Citations	Total link strength
1	Hood (2006)	108	Coffin C., Historical discourse: the language of time, cause and evaluation, (2006)	8	11	Hood (2006)	108	16
2	Oteza (2008)	37	Halliday M.A.K., An introduction to functional grammar, (1994)	9	19	Oteíza (2008)	37	33
3	Whittaker (2020)	32	Hood S., Appraising research: evaluation in academic writing, (2010)	11	16	Whittaker (2020)	32	18
4	Lancaster (2014)	77	Matthiessen C.M.I.M., An introduction to functional grammar, (2004)	6	11	Lancaster (2014)	77	28
5	Szenes (2021)	29	Martin, J. R. & Rose, D., Working with discourse: meaning beyond the clause, (2003)	9	15	Szenes (2021)	29	11
6	Munday (2012)	20	Martin, J. R. & Rose, D, Working with discourse: meaning beyond the clause, (2007)	7	9	Munday (2012)	20	21
7	Miller (2014)	74	Hunston, S., & Thompson, G., Evaluation in text: authorial stance and the construction of discourse, (2000)	7	14	Miller (2014)	74	32
8	Liardét (2018)	33	White P., The language of evaluation: appraisal in english, (2005)	6	7	Liardét (2018)	33	24
9	Llinares (2015)	34	White P.R.R., The language of evaluation: appraisal in english, (2005)	17	28	Llinares (2015)	34	23

4.3. The Co-occurence analysis

The results of the co-occurrence analysis will identify prominent research trends in recent years and, simultaneously, indicate the degree of keyword prominence (density)—that

is, the extent to which particular terms attract scholarly attention and reflect sustained interest within the field.

4.3.1. Dominant Trends

Figure 4 identifies six prominent clusters regarding the application of Appraisal theory to foreign language teaching. Overall, the clustering structure suggests a diversity of implementation scopes and, at the same time, delineates key research directions through central keywords and their discourse-linked networks.

Cluster 1 (yellow) is located in the upper portion of Figure 4, with the central keyword "appraisal" linked to "teacher education," "higher education," and "engagement." This cluster indicates a research trajectory that employs Appraisal theory as an analytical framework to support teacher education, extending to instructional contexts at both undergraduate and postgraduate levels. Notable scholars include Adlington et al. (2024), Godoy-Echiburú and Flores-Calvo (2025), and Nurhadi et al. (2024).

Cluster 2 (green), positioned on the left side of Figure 4, has "academic writing" as the central keyword and connects to "translation," "discourse analysis," "L2 writing," "SFL," and "attitude." The cluster reflects research on teaching academic writing, where discourse analysis is treated as a foundational tool for explaining and guiding text production. Importantly, the outcomes of discourse analysis are also applied to translation studies and L2 writing instruction. Representative works include Munday (2012) and Yu and Wu (2018).

Cluster 3 (light blue) appears on the left, below Figure 4, using "appraisal theory" as the central keyword and showing prominent links to "argumentative essays" and "legal translation." This cluster represents a line of inquiry centered on analyzing texts for specific discourse purposes, particularly argumentative essays and legal translation. Representative authors include Cheung and Low (2017) and Geng et al. (2024).

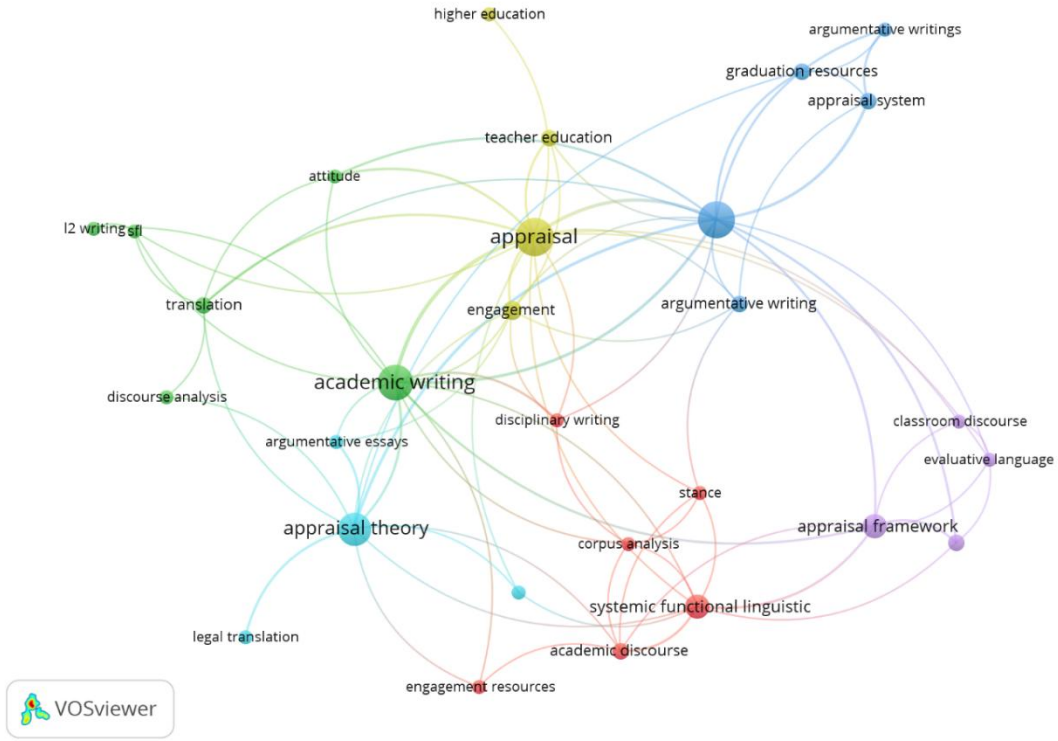
Cluster 4 (red) is located at the bottom of Figure 4, with "Systemic Functional Linguistics" as the central keyword. It is linked to "disciplinary writing," "stance," "corpus analysis," "academic discourse," and "engagement resources." This trend emphasizes investigating interpersonal relations in academic texts and connecting such insights to discipline-specific writing practices. Representative studies include Reznicek-Parrado et al. (2018) and Cederholm (2025).

Cluster 5 (purple) is situated on the right side of Figure 4. With "appraisal framework" as the central keyword, it connects to "classroom discourse and evaluative language." This cluster suggests growing interest in analyzing classroom discourse and evaluative language in educational settings. Representative authors include Cavanna et al. (2015) and Myskow (2018).

Cluster 6 (sky blue) is also located on the right side of Figure 4, taking "argumentative writing" as the central keyword and linking to "graduation resource" and "appraisal system." The principal research direction in this cluster concerns the analysis of graduation in texts intended for argumentative writing. Representative scholars include Yang (2016) and Vo and Le (2025).

Figure 4

Dominant Trends of Appraisal Theory

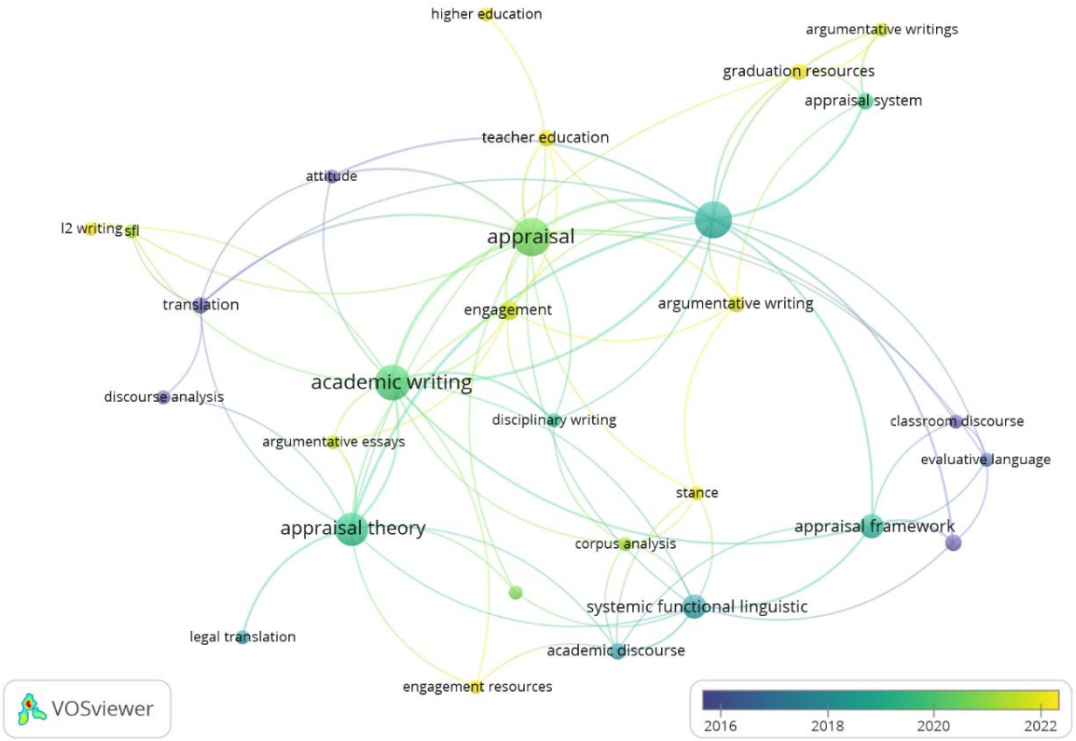


4.3.2. New research trends in the last five years

Figure 5 synthesizes prominent research trends observed over the past five years. The keywords highlighted in bright yellow—L2 writing, higher education, argumentative writings, graduation resources, engagement, and stance—are highly concentrated, indicating that recent studies have largely focused on leveraging appraisal resources, particularly engagement and graduation, to address core issues closely related to writing instruction at the university level. These findings also demonstrate the applicability of appraisal theory in higher education contexts, as evidenced by representative studies such as Cederholm (2025), Hutasuhut et al. (2023), and Ilić (2022).

Figure 5

New research trends in Appraisal Theory over the last five years



4.3.3. Trend Capturing Scholarly Attention

Structurally, within the appraisal theory framework, the keywords highlighted in dark yellow operate as central nodes and maintain direct linkages with the associated research content. Specifically, most studies in this cluster foreground appraisal and appraisal theory. Moreover, among the key terms, academic writing and argumentative writing exhibit comparatively high density, suggesting that these represent the two most distinctly emphasized research directions. This pattern implies that scholars prioritize using appraisal theory to explain and analyze linguistic phenomena related to the construction of academic texts, while also using such insights to inform the design of teaching and learning activities aligned with writing objectives.

Figure 6

Trends Capturing Scholarly Attention in Appraisal Theory



Overall, Figure 6 suggests that the effectiveness of appraisal theory for teaching writing skills has been widely recognized in recent scholarship, as reflected in studies such as Ndoricimpa and Barad (2024) and Li (2024).

4.4. Discussion on the Feasibility of Appraisal Theory in ESL Teaching in Vietnam

In spite of a strong position in linguistics, the application of Appraisal Theory to English as a Language (ELT) teaching in Vietnam remains a potential gap that has not been adequately explored. However, recent studies have been proving the feasibility and high practical value of this theory in complementing the academic competence of Vietnamese learners.

Indeed, Appraisal Theory has indicated its power in contouring 'writers' authorial voice' which is believed to be a key element of academic discourse. Vo an Le (2025) inspected the use of Graduation resources in TESOL master's theses by Vietnamese EFL learners. This research confirms the essential role of instructions in Appraisal Theory for helping students manage professional and persuasive academic voice. Similarly, Chung and Crosthwaite (2024), found significant differences in the expression of attitudes between native English speakers and Vietnamese learners who use English as second language. The fact that the latter often express themselves too directly but lack diversity in their choice of assessment language. This demonstrates that Appraisal Theory is necessary for teachers to guide learners in conveying attitudes subtly and for examiners to more accurately assess academic stances.

Apart from written texts, the elasticity of Appraisal Theory is also displayed in studies of spoken discourse. Ngo and Unsworth (2015) suggested powerful clarifications to the

Attitude system based on group discussion data from graduate students in both English and Vietnamese. These improvements enrich the theoretical framework and at the same time provide valuable learning resources to enhance the effectiveness of expressing students' positions in multilingual contexts.

In addition to the desired results brought about for language learning, especially for teaching English as second language, this theory can be used as a dynamic accompanying framework to evaluate learners' perceptions of technology support in education. Ngan and Le (2025) delivered a study adopting Appraisal Theory. Their paper explores the awareness of 450 students and 12 teachers about the use of Google Translate. The results of this study bring up penetrating perspectives for policymakers on the role of technology. In addition, they confirm that Appraisal Theory is a functional tool capable of comprehensively improving the teaching and learning of foreign languages by understanding attitudes and promoting ethical technology use.

Although the application of Appraisal Theory appeared not to arouse much attention from researchers in Vietnam, the aforementioned studies have soundly established its implication in both general linguistics and teaching English as a Second Language (ESL). In short, the practicality of Appraisal Theory extends a broad range of academic needs—from decoding the complexities of argumentative writing and spoken discourse to evaluating user perceptions of emerging educational technologies. This implication proves that the framework is a descriptive linguistic tool and it also offers a practical, far-reaching methodology for classroom variation. As educators and researchers have had a shift to more communicative and evaluative-based instruction, Appraisal Theory has been becoming a central part in the modernization of English language teaching in Vietnam. Its adoption assures the prospect that learners are able to employ the “evaluative literacy” necessary to navigate the demands of global academic and professional discourse with confidence and precision.

5. Conclusion

In this paper, we mapped the research landscape of applying Appraisal theory in language education by examining publication development, research trends, and the feasibility of this theory within the context of Vietnam. Overall, the field has shown a clear recovery and growth trajectory: after 2020, publication numbers rebounded and continued to rise, reaching a second peak in 2024. In terms of international contributions, Australia leads the field by a wide margin over other countries. At the journal level, *Journal of English for Academic Purposes* and *Linguistics and Education* emerged as the two most prominent venues in this research area. The analysis of authorship impact further indicates that Yadan Li, Ana Llinares, and Ryan T. Miller are key contributors shaping scholarly influence in this domain, while Hood’s (2006) work remains the most cited and has exerted substantial impact on subsequent studies.

Beyond identifying major outlets, authors, and influential prior research, the present study also explored six prominent clusters related to the application of Appraisal theory to foreign language teaching. Importantly, the findings suggest that Appraisal Theory is increasingly becoming a central component in the modernization of English language teaching in Vietnam, offering meaningful theoretical and practical value for advancing how language evaluation, interpersonal meaning, and textual stance can be analyzed and taught in educational settings.

Although bibliometric analysis enables the identification and generalization of research trends regarding the application of Appraisal Theory in foreign language education using

Scopus data, this approach presents several limitations. First, the findings are highly contingent on the selection and frequency of keywords. Consequently, less common keywords—or terms not captured within the principal keyword cluster—may be underrepresented, which can distort estimates of the research community's attention to particular topics (Zahra et al., 2021). In addition, newly published studies, especially exploratory work or research that has not yet attracted substantial citation uptake, may not be fully indexed in the database, resulting in gaps within the analytical dataset (Thelwall, 2018). These constraints are partly attributable to the indexing characteristics of bibliographic databases, particularly open-access data infrastructures.

Second, restricting the analysis to a single database (e.g., Scopus) may limit the comprehensiveness of the study, as relevant publications are not always indexed within the same system. To mitigate these issues, future research should integrate multiple bibliographic sources and adopt complementary analytical methods. For example, time-series trend analysis and author-based citation network construction could be used alongside bibliometric indicators to provide a more nuanced and robust account of how the field has evolved over time. Collectively, these strategies would improve both the completeness and accuracy of the evidence regarding the development of this research area.

References

Adlington, R., O'Neill, K., Volpe, C. R., & Harrington, I. (2024). Promoting a sense of belonging in university online learning: How and why initial teacher education students experience an increased sense of belonging. *Australasian Journal of Educational Technology*, 40(5), 84. https://doi.org/10.14742/ajet.9487

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2), 191–235. https://doi.org/10.1177/053901883022002003

Caputo, A., Fiorentino, R., & Garzella, S. (2019). From the boundaries of management to the management of boundaries: Business processes, capabilities and negotiations. *Business Process Management Journal*, 25(3), 391–413. https://doi.org/10.1108/BPMJ-11-2017-0334

Cavanna, J. M., Herbel-Eisenmann, B., & Seah, W. T. (2015). Investigating teachers' appraisal of unexpected moments and underlying values: An exploratory case in the context of changing mathematics classroom discourse. *Research in Mathematics Education*, 17(3), 163–182. https://doi.org/10.1080/14794802.2015.1112301

Cederholm, P. (2025). Explain and engage: linguistic resources in swedish as a second language textbooks to support academic writing. *Language and Education*, 40(2), 329-346. https://doi.org/10.1080/09500782.2025.2521432

Cheung, Y. L., & Low, T. H. (2017). Pre-university students' voice construction in argumentative essays. *RELC Journal*, 50(2), 269–284. https://doi.org/10.1177/0033688217716508

Chung, T. T., & Crosthwaite, P. R. (2024). Stance-taking through APPRAISAL in L1 and L2 English argumentative essays: Insights from Vietnamese L2 English. *IRAL - International Review of Applied Linguistics in Language Teaching*, 64(1), 265–299. https://doi.org/10.1515/iral-2024-0079

Geng, Y., Chen, G., & Li, M. (2024). Analyzing engagement strategies in argument chain: A comparison between high- and low-scoring EFL undergraduate argumentative essays. *Journal of English for Academic Purposes*, 71, Article 101428. https://doi.org/10.1016/j.jeap.2024.101428

Godoy-Echiburú, G., & Flores-Calvo, L. (2025). Construcción de la dialogicidad en trabajos escritos de estudiantes del profesorado de Literatura en Uruguay. *Logos: Revista de Lingüística, Filosofía y Literatura*, 35(2), 844–867. https://doi.org/10.15443/RL3550

Halliday, M. A. K. (1994). *An introduction to functional grammar*. London: Edward Arnold.

Hood, S. (2006). The persuasive power of prosodies: Radiating values in academic writing. *Journal of English for Academic Purposes*, 5(1), 37–49. https://doi.org/10.1016/j.jeap.2005.11.001

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford University Press. https://doi.org/10.1093/oso/9780198238546.001.0001

Hutasuhut, M. L., Chen, H., & Matruglio, E. (2023). Engaged at the first sight! Anticipating your audience as a way to think critically in writing an argument. *Indonesian Journal of Applied Linguistics*, 12(3), 694-707. https://doi.org/10.17509/ijal.v12i3.55170

Ilić, K. M. (2022). Engagement resources in esp research article abstracts. *PHILOLOGIST – Journal of Language, Literature, and Cultural Studies*, 13(25), 240–264. https://doi.org/10.21618/fil2225240i

Iwami, S., Ojala, A., Watanabe, C., & Neittaanmäki, P. (2020). A bibliometric approach to finding fields that co-evolved with information technology. *Scientometrics*, 122(1), 3–21. https://doi.org/10.1007/s11192-019-03284-9

Lancaster, Z. (2014). Exploring valued patterns of stance in upper-level student writing in the disciplines. *Written Communication*, 31(1), 27–57. https://doi.org/10.1177/0741088313515170

Li, R. (2024). Beyond surfaces and depths: An SFL analysis of fine gradations of meaning in undergraduates' writing about literature. *Written Communication*, 42(1), 85–119. https://doi.org/10.1177/07410883241286899

Martin, J. R., & White, P. R. (2005). *The Language of Evaluation: Appraisal in English*. Basingstoke: Palgrave Macmillan.

Martin, J. R., & Rose, D. (2003*). Working with discourse: Meaning beyond the clause*. Continuum.

Martin, J. R., & Rose, D. (2007). *Working with Discourse: Meaning Beyond the Clause (2nd ed.)*. Continuum International Publishing Group.

Miller, R. T., Mitchell, T. D., & Pessoa, S. (2014). Valued voices: Students' use of engagement in argumentative history writing. *Linguistics and Education*, 28, 107–120. https://doi.org/10.1016/j.linged.2014.10.002

Munday, J. (2012). New directions in discourse analysis for translation: A study of decision-making in crowdsourced subtitles of Obama's 2012 State of the Union speech. *Language and Intercultural Communication*, 12(4), 321–334. https://doi.org/10.1080/14708477.2012.722099

Myskow, G. (2018). Changes in attitude: Evaluative language in secondary school and university history textbooks. *Linguistics and Education*, 43, 53–63. https://doi.org/10.1016/j.linged.2017.12.001

Ndoricimpa, C., & Barad, D. P. (2024). Writing in the disciplines: Comparison of critical stance qualities in literary analyses by university students writing in L1 and L2 contexts. *Indonesian Journal of Applied Linguistics*, 14(2), 415–428. https://doi.org/10.17509/ijal.v14i2.74890

Ngan, C. H., & Le, T. H. V. (2025). Vietnamese EFL teachers' and students' perceptions of using Google Translate in the classroom. *Computer-Assisted Language Learning Electronic Journal*, 26(3). 228-247. https://doi.org/10.54855/callej.252638

Ngo, T., & Unsworth, L. (2015). Reworking the appraisal framework in ESL research: Refining attitude resources. *Functional Linguistics*, 2(1), Article 1. https://doi.org/10.1186/s40554-015-0013-x

Nurhadi, K., Rahman, A., Khasyar, M. L., & Suharwanto, S. (2024). The emotional geography of faculty members engaging in virtual teacher professional development (VTPD):

evidence from indonesia. Journal of Applied Research in Higher Education, 17(2), 663-674. https://doi.org/10.1108/jarhe-06-2022-0187

Oteíza, T. (2017). The appraisal framework and discourse analysis. In G. O. Tom Bartlett, *The Routledge handbook of systemic functional linguistics* (pg. 457-472). Abingdon: Routledge.

Phan, T. L. (2022). Lí thuyết đánh giá trong phân tích diễn ngôn: nghiên cứu thử nghiệm từ cáo trạng tiếng việt. *Tạp chí khoa học trường Đại học Sư phạm TP Hồ Chí Minh*, 19(12), 2146-2159.

Reznicek-Parrado, L. M., Patiño-Vega, M., & Colombi, M. C. (2018). Academic peer tutors and academic biliteracy development in students of Spanish as a heritage language. *Journal of Spanish Language Teaching*, 5(2), 152–167. https://doi.org/10.1080/23247797.2018.1538358

Szenes, E., & Tilakaratna, N. (2021). Deconstructing critical reflection in social work and business: Negotiating emotions and opinions in reflective writing. *Journal of English for Academic Purposes*, 49, Article 100931. https://doi.org/10.1016/j.jeap.2020.100931

Thelwall, M. (2018). Dimensions: a competitor to scopus and the web of science? *Journal of Informetrics*, 12(2), 430-435. https://doi.org/10.1016/j.joi.2018.03.006

van Eck, N. J., & Waltman, L. (2007). VOS: A new method for visualizing similarities between objects. In R. Decker & H.-J. Lenz (Eds.), *Advances in data analysis* (pp. 299–306). Springer. https://doi.org/10.1007/978-3-540-70981-7_34

Vo, L.-H., & Le, L. V. (2025). Exploring EFL Writers' Authorial Voice in Academic Discussions: A Graduation-Based Appraisal Analysis. *Forum for Linguistic Studies*, 7(5), 932–946. https://doi.org/10.30564/fls.v7i5.8873

Wei, Y., Wherrity, M., & Zhang, Y. (2015). An Analysis of Current Research on the Appraisal Theory. *Linguistics and Literature Studies*, 3(5), 235–239.

Whittaker, R., & McCabe, A. (2020). Expressing evaluation across disciplines in primary and secondary CLIL writing: a longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 26(3), 345-362. https://doi.org/10.1080/13670050.2020.1798869

Yang, Y. (2016). Appraisal resources in Chinese college students' English argumentative writing. *Journal of Language Teaching and Research*, 7(5), 1002–1013. https://doi.org/10.17507/jltr.0705.23

Yu, H., & Wu, C. (2018). Attitude as mediation: Peritextual commentary in the translation of the Platform Sutra. *Text & Talk*, 38(5), 633–654. https://doi.org/10.1515/text-2018-0017

Zahra, A. A., Nurmandi, A., Tenario, C. B., Rahayu, R., Benectitos, S. H., Mina, F. L. P., & Haictin, K. M. (2021). Bibliometric analysis of trends in theory-related policy publications. *Emerging Science Journal*, 5(1), 96-110. https://doi.org/10.28991/esj-2021-01261

Bionote

Ly Tuan Phan earned a PhD in Contrastive Linguistics from the University of Social Sciences and Humanities, Ho Chi Minh City - Vietnam National University. He is currently a lecturer at Ho Chi Minh City University of Law. His research interests include legal discourse and Systemic Functional Linguistics.

Truc Ngoc Thanh Hoang holds an MA in Linguistics and is currently a lecturer at Ho Chi Minh City University of Law. With a focus on the intersection of linguistics and law, her scholarly interests include legal discourse analysis and the application of linguistic theories to the study and teaching.

Dung Thi Thuy Pham holds an MA in TESOL and is currently a PhD candidate in Contrastive Linguistics. With a background in English language pedagogy, her research has evolved to specialize in the intersection of language and law. Her current scholarly work focuses on legal English teaching and the linguistic nuances of legal discourse. Looking forward, her research interests involve contrastive linguistics within legal frameworks, specifically examining the structural and functional disparities between legal languages.

#56 Exploring Undergraduate Students' Use of AI for Academic Presentations in Higher Education

Nguyễn Thị Hồng Liên[1], Nguyễn Thị Thanh Vân[2]

Abstract

This study examines undergraduate students' self-reported use of generative artificial intelligence (GenAI) for academic presentation design and their evaluations of its usefulness and ease of use. Survey data were collected from 197 undergraduate students at a Vietnamese university. The study used descriptive statistics, multiple-response analysis, reliability testing, factor-based measurement validation, and K-means clustering to analyze GenAI usage patterns, task delegation, and student perceptions. The findings show that GenAI had become a routine support tool, with 86.8% of students reporting weekly or daily use. However, frequent use did not necessarily indicate full dependence, as most students reported moderate task delegation. Students used GenAI most often for content preparation and less often for live presentation tasks, suggesting an asymmetric pattern of delegation across the presentation workflow. Three user profiles emerged: Power Integrators, Selective Minimalists, and Skeptical Delegators. These profiles show that students differed not only in usage frequency, but also in tool diversity, task delegation, perceived autonomy, and attitudes toward GenAI. Students evaluated GenAI positively in terms of perceived usefulness and ease of use, but valued it more for immediate productivity than for long-term presentation skill development. The study highlights the need for pedagogically guided GenAI integration, particularly in relation to task delegation, output evaluation, disclosure, and reflective revision.

***Keywords*:** GenAI, academic presentation, higher education, student perception

1. Introduction

Generative artificial intelligence (GenAI) has rapidly transformed the landscape of higher education, evolving from a technological novelty into an embedded educational resource (Lim et al., 2023). Following the public release of large language models like ChatGPT, students have gained unprecedented access to tools that can generate human-like text, synthesize complex information, and assist in multifaceted academic tasks (Dwivedi et al., 2023). This rapid proliferation has shifted institutional discourses from initial attempts at prohibition to a growing recognition of the need for AI literacy, ethical integration, and human-centered pedagogical strategies (UNESCO, 2023). Consequently, GenAI is no longer viewed merely as an external disruption but as a permanent fixture in the academic ecosystem that demands rigorous empirical investigation.

Understanding how undergraduate students utilize GenAI for academic work is particularly crucial. As the primary early adopters of these technologies, university students—especially those belonging to Generation Z—exhibit a strong willingness to employ GenAI to enhance their productivity, digital competence, and learning efficiency (Chan & Lee, 2023). Student perceptions and interactions with their learning environments directly influence their educational approaches and eventual outcomes (Chan & Hu, 2023). While AI systems offer significant advantages, such as providing personalized feedback, facilitating brainstorming, and overcoming language barriers, they also pose profound challenges. Educators frequently

[1] *Corresponding author, MA in TESOL, Faculty of Foreign Languages, HUFLIT, Ho Chi Minh City; Vietnam. Email: liennth1@huflit.edu.vn*
[2] *MA in TESOL, Dong Nai University, Vietnam; Emai: ms.van0708@gmail.com*

express concerns regarding over-reliance, diminished critical thinking, and the potential illusion of comprehension when students outsource cognitive labor to algorithms (Kim et al., 2024; Sullivan et al., 2023). Therefore, empirical investigations into students' reported usage patterns and AI-supported academic practices are essential.

Within the spectrum of academic assignments, academic presentations represent a uniquely demanding, multimodal activity. Developing an effective presentation requires students to synthesize rigorous research, conceptual organization, visual design, and oral delivery (Chu et al., 2025). The emergence of GenAI has introduced new mechanisms for cognitive offloading in this space, enabling students to utilize AI-assisted tools for organizing information, outlining narratives, and formatting visual slides (Georgiev & Tinsley, 2024). However, despite the inherent complexity of presentation design and the increasing availability of both text-based and multimodal GenAI tools, this specific domain remains an underexplored area of student AI use.

Current research has not fully captured how students report using GenAI across different academic tasks. Most studies only show general statistics about GenAI use for academic writing, rather than examining how students use it for different parts of complex tasks like presentations or how usage differs among different groups of students. Furthermore, while the Technology Acceptance Model (TAM) has been widely applied to assess educational technologies through its core constructs of Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) (Davis, 1989; Venkatesh et al., 2003), its application to GenAI requires further scrutiny. For conversational AI tools whose primary interface is natural language, the traditional boundary between a tool's utility and its operational effortlessness may become increasingly blurred (Strzelecki, 2023). It remains unclear whether PU and PEOU retain their conceptual distinctness in design-oriented contexts, highlighting a significant gap in understanding how students genuinely evaluate these tools.

To address these theoretical and empirical gaps, the present study aims to explore undergraduate students' reported usage patterns of GenAI for academic presentation design. By identifying detailed usage patterns, examining the extent of task delegation across different presentation stages, and analyzing students' evaluations of these tools, this research provides essential insights into the intersection of AI adoption and academic design tasks. Specifically, this study is guided by the following two research questions:

- **RQ1.** How do undergraduate students use GenAI for academic presentation design, particularly in terms of usage patterns and task delegation?
- **RQ2.** How do students evaluate GenAI tools for academic presentation design in terms of perceived usefulness and perceived ease of use?

2. Literature Review

2.1. Integration and Usage Patterns of Generative AI in Higher Education

The integration of Generative Artificial Intelligence (GenAI) is catalyzing a paradigm shift in educational practices, moving institutional focus from strict prohibition to the development of AI literacy and critical interaction skills (UNESCO, 2023). GenAI has rapidly transitioned from a technological novelty to an embedded educational tool, fundamentally changing the landscape of higher education (Lim et al., 2023). Students have quickly adopted tools such as ChatGPT for a wide array of academic tasks, recognizing their capacity to provide personalized, real-time support (Chan & Hu, 2023). A recurring theme in the literature is the conceptualization of GenAI as a versatile, multi-tasking assistant. Students frequently use these

tools to brainstorm ideas, retrieve information, and refine their writing, effectively treating the AI as a virtual tutor or a digital peer during their academic workflows (Kim et al., 2024). Furthermore, research highlights a distinct generational enthusiasm, with Generation Z students demonstrating a strong willingness to integrate GenAI to enhance their productivity, digital competence, and overall academic efficiency (Chan & Lee, 2023). Students are increasingly positioned as active users who interact with GenAI to test ideas, explore perspectives, and support academic meaning making (Kim et al., 2024). However, research also shows that students do not all use GenAI in the same way. How much and how well they use it often depends on their previous experience with technology, their motivation to learn, and their field of study (Chu et al., 2025; Li et al., 2026). While some learners actively exploit the multimodal capabilities of GenAI to explore creative problem-solving and conceptual abstraction, others limit their engagement to superficial tasks like grammar correction and basic translation (Chu et al., 2025; Ngo, 2023). Because of these differences, researchers should not only look at overall rates of AI use. They also need to examine how different groups of students use GenAI in specific ways for their academic work.

2.2. Task Delegation in Multimodal Academic Presentations

Academic presentations represent a uniquely demanding educational activity that requires students to synthesize rigorous research, coherent textual organization, and engaging visual design (Georgiev & Tinsley, 2024). The emergence of GenAI has introduced new possibilities for cognitive offloading and task delegation in this multimodal context. Creating presentations involves the cognitive challenge of harmonizing textual arguments with visual representations and oral delivery. Research indicates that utilizing GenAI to manage these multimodal demands provides a supportive and inclusive environment, enabling students to express their ideas innovatively while maintaining their learning motivation (Chu et al., 2025). Studies indicate that students are increasingly leveraging GenAI, such as Microsoft PowerPoint Designer or ChatGPT, to organize information, format slides, and outline their academic presentations (Baker et al., 2024; Georgiev & Tinsley, 2024). These tools are praised for their ability to streamline the preparatory phases of presentation design, allowing students to save time and reduce the cognitive burden associated with visual structuring (Baker et al., 2024; Li et al., 2026).

Despite these affordances, research shows that students differ in how much control they are willing to give AI. While GenAI is frequently tasked with early-stage preparation—such as generating subtitles, synthesizing reading materials, and proposing structural outlines—students often exhibit caution when it comes to delegating the final design and performance elements (Georgiev & Tinsley, 2024; Kim et al., 2024). This reluctance stems from well-documented concerns about AI's limitations in contextual understanding, logical coherence, and the alignment of visual elements with textual meaning (Georgiev & Tinsley, 2024). Moreover, studies on specific populations, such as pharmacy students preparing didactic patient presentations, reveal that while GenAI effectively organizes clinical information and improves student confidence, its success heavily depends on the user's prior experience and domain knowledge (Baker et al., 2024). Furthermore, students report anxieties regarding the potential loss of their authentic voice and the dilution of human emotion in their presentations, which are critical for effective audience engagement (Chu et al., 2025; Kim et al., 2024). Consequently, a shared finding across studies is that while students readily delegate routine preparatory tasks to GenAI, they strongly prefer to retain human oversight and control over the final presentation to ensure authenticity and logical rigor.

2.3. Student Evaluations of GenAI through the Technology Acceptance Model

To systematically understand students' adoption of GenAI, researchers have frequently relied on the Technology Acceptance Model (TAM), originally developed by Davis (1989), which posits Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) as the primary drivers of technology acceptance (King & He, 2006). In the context of GenAI, studies highlight that personal innovativeness, habit, and hedonic motivation significantly predict students' behavioral intention to utilize tools like ChatGPT (Strzelecki, 2023). The conversational and intuitive interfaces result in exceptionally high PEOU evaluations, as students find the barrier to entry remarkably low (Ngo, 2023). The ease with which students can enter prompts, adjust outputs, and interact with the AI significantly bolsters their behavioral intention to continue using these technologies for academic tasks (Strzelecki, 2023).

Conversely, student evaluations of Perceived Usefulness (PU) present a more paradoxical picture. On one hand, students highly rate the immediate benefits of GenAI, such as rapid task completion, enhanced writing quality, and the provision of immediate, personalized feedback (Chan & Hu, 2023; Kim et al., 2024). On the other hand, this perception of usefulness is frequently counterbalanced by profound concerns over academic integrity, algorithmic bias, and the reliability of AI-generated content (Dwivedi et al., 2023; Perkins, 2023). Students are especially aware of the phenomenon of AI "hallucinations"—the generation of plausible but factually incorrect information and fabricated citations—which directly threatens the rigorous standards required in higher education (Kim et al., 2024; Ngo, 2023). Furthermore, a significant debate exists regarding the long-term impact of GenAI on holistic competency development. While students appreciate the short-term productivity gains, they simultaneously express concern that an over-reliance on AI may hinder their independent critical thinking, problem-solving skills, and social interaction skills (Chan & Hu, 2023; Chu et al., 2025). This suggests that students view GenAI in a practical way. Students appear to value GenAI primarily for its immediate contribution to task efficiency, while remaining less certain about its role in deeper competence development.

While the existing literature offers a strong foundation for understanding the general adoption, perceived benefits, and ethical concerns surrounding GenAI in higher education, an important research gap remains. Much of the current literature has focused on traditional academic writing or broad patterns of use, rather than examining how students allocate specific tasks to GenAI within complex, multimodal assignments such as academic presentations. Furthermore, empirical data exploring the relationship of perceived usefulness and ease of use in these specific, design-oriented contexts is limited. The present study addresses this identified gap by systematically examining undergraduate students' self-reported use of GenAI for academic presentation design. By investigating their nuanced usage patterns, specific task delegation practices, and detailed evaluations of GenAI tools, this study contributes essential empirical insights into how students navigate the intersection of human creativity and artificial intelligence in modern higher education.

3. Methodology

3.1 Participants and Procedure

Participants were 197 undergraduate students enrolled at a Vietnamese university in Ho Chi Minh City during the 2025–2026 academic year. The sample comprised 131 female (66.5%) and 65 male (33.0%) students, with one participant identifying as they/them (0.5%); one response was excluded from the initial 198 due to an uninterpretable entry. Participants were predominantly third-year students (n = 127, 64.5%), with second-year (n = 60, 30.5%) and fourth-year-or-above (n = 10, 5.1%) also represented. Data were collected via a Google

Forms survey using the convenience sampling method. Participation was voluntary and anonymous; no personally identifying information was collected.

3.2. Instrument

The study collected data through a survey that consisted of four sections. The first collected demographic and background information, including year of study, gender, prior AI experience, and frequency of AI use for presentation preparation. The second examined usage patterns, covering the purposes of AI use, types of AI tools employed, the proportion of work delegated to AI, and disclosure practices. The third explored students' attitudes, including their perceived ability to complete presentations without AI, awareness of citation and attribution norms for AI-assisted work, and perceived changes in presentation skills since adopting AI. The fourth section comprised 5-point Likert-scale items based on the Technology Acceptance Model (TAM) and its subsequent extensions. Adapted from the foundational work of Davis (1989), the TAM2 framework by Venkatesh and Davis (2000), and the Unified Theory by Venkatesh et al. (2003), this section was contextualized for GenAI. It measured three core constructs: Perceived Usefulness, Perceived Ease of Use, and Behavioral Intention to continue using GenAI for future academic presentations.

3.3. Measurement Validation

All analyses were conducted using IBM SPSS Statistics. Before addressing the research questions, the three TAM-based constructs were examined to confirm that the survey items were appropriate and reliable. The results showed that the items were suitable for factor analysis and formed a clear three-factor structure, explaining 71.4% of the total variance. The three scales also demonstrated good reliability: Perceived Usefulness ($\alpha = .79$, $M = 3.94$, $SD = .67$), Perceived Ease of Use ($\alpha = .91$, $M = 4.00$, $SD = .74$), and Behavioral Intention ($\alpha = .91$, $M = 4.14$, $SD = .82$). In addition, the data met the basic assumptions for further statistical analysis. Composite scores were then calculated by averaging the items within each construct, and these scores were used in the subsequent analyses.

4. Results

4.1 Research Question 1: Usage Patterns and Task Delegation

RQ1 examined the current usage patterns and levels of task delegation when undergraduate students employ GenAI for presentation design. This question was addressed through three analytic steps: (a) descriptive analysis of usage intensity, (b) multiple-response analysis of purposes and tools, and (c) K-means cluster analysis to identify user typologies. Following established methodological precedents in educational AI research (Luckin et al., 2022), this clustering approach was used to segment students into three distinct behavioral profiles—Power Integrators, Selective Minimalists, and Skeptical Delegators—facilitating a targeted analysis of their task delegation strategies.

4.1.1 Intensity and Modes of GenAI Use

Descriptive analyses revealed that GenAI has become a routine component of presentation preparation rather than an occasional aid. As Table 1 shows, **86.8% of participants reported using GenAI at least weekly**, with 43.7% using it daily. Only 4.1% reported rare or no use. This frequency distribution meaningfully exceeds the adoption levels reported in previous empirical studies (e.g., Chan & Hu, 2023; Strzelecki, 2023), suggesting that within this sample, GenAI use was highly common.

Table 1

Frequency Distribution of Usage Intensity, Delegation, and Perceived Outcomes (N = 197)

Variable	Response option	n	%
Usage frequency	Never / rarely	8	4.1
	Monthly (1–3×/month)	18	9.1
	Weekly (1–3×/week)	85	43.1
	Daily (≥ 1×/day)	86	43.7
Delegation (% workload by AI)	0–25%	26	13.2
	26–50%	116	58.9
	51–75%	44	22.3
	76–100%	11	5.6
Disclosure of AI use	Yes	92	46.7
	No	27	13.7
	Have not considered	78	39.6
Self-rated skill change	Significantly decreased	1	0.5
	Slightly decreased	4	2.0
	No change	50	25.4
	Increased slightly	96	48.7
	Increased significantly	46	23.4

Although students reported frequent use of GenAI, the data suggest that they used it mainly as a support tool rather than as a replacement for their own work. Most students (58.9%) said GenAI contributed between 26% and 50% of their presentation workload, while heavy delegation was uncommon, with only 5.6% reporting AI involvement at the 76–100% level. Minimal delegation was also relatively low at 13.2%. This pattern challenges the common concern about widespread "AI over-reliance" in education, as students appeared to retain substantial control over the final product. However, their disclosure practices were less consistent. While 46.7% reported openly acknowledging AI use, only 13.7% reported not disclosing it, and a large proportion, 39.6%, had not considered disclosure at all. This suggests that many students may lack clear guidance on AI attribution, pointing to a need for stronger institutional policy and curriculum support (Perkins, 2023). At the same time, students generally viewed GenAI as beneficial for their development, with 72.1% reporting slight or significant improvement in presentation skills and only 2.5% reporting decline. Overall, the findings present a balanced picture: students used GenAI frequently, but mostly in a controlled and supportive way, while issues of disclosure and academic guidance remain underdeveloped.

4.1.2 Purposes and Tools: An Asymmetric Delegation Pattern

Multiple-response analysis revealed a distinctive asymmetry in how students distribute AI assistance across the presentation workflow (Table 2). Preparation-stage tasks were extensively delegated: 93.4% of participants used AI for content preparation. In contrast, performance-stage tasks remained largely human-driven: only 34.0% used AI during the live presentation itself. Design (49.2%) and practice/feedback (41.6%) occupied intermediate positions in this gradient.

Table 2

Multiple-Response Frequencies for Purposes and Tools (N = 197)

Category	Option	n responses	% of responses	% of cases
Purposes of GenAI use	Content preparation	184	42.8	93.4
	Presentation design	97	22.6	49.2
	Practice / feedback	82	19.1	41.6
	Live presentation	67	15.6	34.0
Tools used	Content generators	191	47.6	97.0
	AI search	103	25.7	52.3
	Slide creators	68	17.0	34.5
	Image generators	21	5.2	10.7
	Video generators	17	4.2	8.6
	Other	1	0.2	0.5

Note. Percentages of cases sum to more than 100 because participants could select multiple options.

Students showed a clear difference in how they used AI across presentation tasks. They seemed more willing to use AI for preparation work, such as generating ideas, organizing content, and drafting text, but they kept more control over the actual performance stage, such as speaking in front of an audience. This pattern aligns with recent findings that students prefer to use AI for structural preparation while retaining human oversight and control over final presentation outcomes (Georgiev & Tinsley, 2024). It also gives a more balanced view than the common concern about AI over-reliance. Rather than using AI for everything, students appeared to make careful choices about when AI support was useful and when they needed to rely on themselves.

A similar pattern appeared in the types of tools students used. Text-based GenAI tools were the most common, with 97.0% using content generators such as ChatGPT, Gemini, or DeepSeek, and 52.3% using AI-powered search tools. In contrast, fewer students used multimodal tools: 34.5% used slide creation tools such as Gamma or Tome, 10.7% used image generators such as Midjourney or DALL-E, and 8.6% used video generators such as Synthesia or Runway. This suggests that students did not simply adopt every kind of AI tool available. Instead, their use followed a text-first pattern, probably because academic presentations depend heavily on research, writing, and script preparation. Practical issues such as cost, difficulty of use, and unclear relevance may also explain why multimodal tools were less widely used.

4.1.3 User Typology: Three GenAI Adopter Profiles

Although the descriptive results showed that GenAI adoption was generally high, they did not fully show the differences in how students used these tools. To examine these differences more clearly, a K-means cluster analysis was conducted using six standardized indicators: usage frequency, delegation percentage, number of presentation stages using AI, number of AI tool categories used, perceived dependence on AI, and perceived usefulness. The analysis found that a three-cluster solution was the clearest and most balanced. All six variables showed significant differences across the clusters, with F values ranging from 21.34 to 59.96, all $p < .001$, and η^2 values from .18 to .38, indicating medium-to-large effects. Other options, including four-cluster and TwoStep solutions, were also tested, but they were either less balanced or less meaningful in explaining the data.

The three clusters differed systematically across behavioral and attitudinal indicators, as shown in Table 3. They were interpreted as distinct adopter profiles and labeled accordingly: *Power Integrators*, *Selective Minimalists*, and *Skeptical Delegators*.

Table 3

Profile of the Three GenAI User Clusters (N = 197)

Indicator	Cluster 1: Power Integrators (n = 84, 42.6%)	Cluster 2: Selective Minimalists (n = 72, 36.5%)	Cluster 3: Skeptical Delegators (n = 41, 20.8%)	F(2, 194)	p
Usage frequency (1–5)	4.65	3.99	3.93	21.34	< .001
Delegation band (1–4)	2.54	1.79	2.24	24.79	< .001
N presentation stages (0–4)	2.73	1.49	2.29	42.60	< .001
N tool categories (0–6)	2.80	1.33	1.71	50.60	< .001
Can complete without AI (1–5)	3.11	3.72	2.51	30.17	< .001
Perceived Usefulness (1–5)	4.32	3.90	3.19	59.96	< .001
Perceived Ease of Use (1–5)	4.28	3.95	3.51	18.76	< .001
Intention to continue use (1–5)	4.55	4.01	3.54	28.28	< .001

The cluster results suggest that students did not use GenAI in one single way. Instead, their use formed three clear patterns, ranging from broad and confident integration to more limited or uneasy use. The largest group was the Power Integrators (n = 84, 42.6%). These students showed the strongest level of GenAI adoption. They used GenAI most frequently (M = 4.65), almost at the level of daily use, and applied it across more presentation stages than the other groups (M = 2.73 out of 4). They also used the widest range of tools (M = 2.80 out of 6). What makes this group stand out is their use of multimodal tools. Compared with the full sample, they were much more likely to use slide-creation tools (63% vs. 34.5%), AI search tools (81% vs. 52.3%), and image generators (19% vs. 10.7%). Their attitudes were also the most positive, with the highest scores for perceived usefulness (M = 4.32), perceived ease of use (M = 4.28), and intention to continue using GenAI (M = 4.55).

A second group, the Selective Minimalists (n = 72, 36.5%), used GenAI in a much more careful and limited way. They were not avoiding AI, but they did not spread their use across many tools or tasks. Most of them used it for content preparation, with 88% reporting this purpose. However, only 7% used slide creators, and only 28% used AI search tools. This group also reported the strongest sense of independence, with the highest score for being able to complete a presentation without AI (M = 3.72). Their workload delegation was the lowest (M = 1.79), mainly falling within the 0–25% range. Even so, their views of GenAI were still generally positive, as shown by their perceived usefulness score (M = 3.90) and intention to continue using it (M = 4.01). This suggests that they were not skeptical users. Rather, they seemed to set boundaries around AI use and kept it mainly as a support tool.

The smallest group, the Skeptical Delegators (n = 41, 20.8%), showed the most complicated pattern. These students seemed to rely on GenAI, but they did not evaluate it as positively as the other groups. They had the lowest autonomy score (M = 2.51), suggesting that they found it difficult to complete presentations without AI. At the same time, they also had the lowest scores for perceived usefulness (M = 3.19), perceived ease of use (M = 3.51), and intention to continue using GenAI (M = 3.54). This creates a somewhat uneasy profile: they appear to need AI, but they are not fully comfortable with it or convinced of its value. For

teaching practice, this group may be especially important because they may need more guidance on how to use GenAI with clearer purpose, less dependence, and greater confidence.

4.2 Research Question 2: Students' Perceptions of GenAI Usefulness and Ease of Use

RQ2 examined how students perceive the usefulness and ease of use of GenAI tools for academic presentations. At the construct level, students showed generally positive views of both TAM dimensions. The mean score for Perceived Usefulness was relatively high, PU *(M = 3.94, SD = .67)*, and the mean score for Perceived Ease of Use was also high, PEOU *(M = 4.00, SD = .74)*. Since both scores were above the scale midpoint of 3.00, the results suggest that students generally found GenAI useful and easy to use. However, the scores were not extremely high, which means students' views were positive but not overly idealized. In other words, they appeared to recognize the value of GenAI while still leaving room for some caution or critical judgment. A closer look at the individual items provides a more detailed picture of how students understood GenAI's usefulness and ease of use.

Table 4

Item-Level Descriptive Statistics for PU and PEOU (N = 197)

Construct	Item	Item content	*M*	*SD*
PU	PU1	AI helps me complete presentations faster	**4.29**	.82
	PU2	AI improves content / idea quality	4.25	.82
	PU4	I am satisfied with AI suggestions / content	3.68	.85
	PU6	My presentation skills have clearly improved with AI	**3.54**	.95
PEOU	PEOU1	Easy to learn how to interact with AI	3.94	.78
	PEOU2	Entering prompts and adjusting outputs is not complex	4.06	.91
	PEOU4	Using AI is intuitive and quick	3.94	.78
	PEOU5	I easily understand and adjust AI-generated content	4.08	.87

The PU items suggest what this study calls a productivity-development asymmetry. Students gave the strongest support to items related to immediate task completion. For example, PU1, "AI helps me complete presentations faster," had the highest mean score *(M = 4.29)*, followed closely by PU2, "AI improves content / idea quality" *(M = 4.25)*. In contrast, the item related to longer-term skill growth, PU6, "my presentation skills have clearly improved," received the lowest mean score *(M = 3.54)*. The difference between PU1 and PU6 was 0.75 points, which is meaningful on a 5-point scale. This suggests that students mainly understand GenAI's usefulness in terms of saving time and improving the quality of the immediate product, rather than helping them gradually develop their own presentation skills. In practical terms, students seem to value what AI helps them produce more than what AI helps them learn. This distinction is important for how teachers design and explain AI-assisted assignments, as discussed in Section 5.

The PEOU items showed a more even pattern, with mean scores ranging from 3.94 to 4.08. This means that students generally found GenAI easy to use across different stages, including learning how to use it, writing prompts, interacting with the tool, and editing its output. The highest-rated item was PEOU5, "I easily understand and adjust AI-generated content" *(M = 4.08)*. This suggests that students felt confident not only in using GenAI, but also in reviewing and modifying what it produced. Therefore, based on their self-reported ease of use, students did not see themselves as simply accepting AI-generated content passively.

Rather, they appeared to see themselves as active editors who could adapt AI output to fit their own presentation needs.

5. Discussion

5.1 Adoption Intensity

The finding that 86.8% of students used GenAI at least weekly, with 43.7% using it daily, suggests that GenAI has become a routine part of academic presentation preparation in this sample. However, frequent use did not necessarily mean full dependence. Most students reported that AI contributed only 26–50% of their presentation workload, while very high delegation remained uncommon. This indicates that students generally used GenAI as a support tool rather than as a complete substitute for their own work.

At the same time, the findings show a gap between adoption and responsible-use awareness. Although 46.7% of students reported disclosing AI use, 39.6% had not considered disclosure. This suggests that students may be using GenAI more quickly than they are developing clear understandings of attribution and academic integrity. Therefore, the main pedagogical issue is no longer whether students use GenAI, but how well they are guided to use it transparently, critically, and appropriately.

5.2. The User Typology

The identification of three distinct user profiles including Power Integrators, Selective Minimalists, and Skeptical Delegators demonstrates that students are not a uniform group. This suggests that the quality of engagement can vary independently from the quantity of use, a pattern that would be missed if the study only reported overall adoption rates. The differences among these groups are particularly revealing when examining how students balance AI assistance with their own academic autonomy. For instance, the Skeptical Delegators cluster highlights a particularly vulnerable subgroup of learners. These students report lower autonomy in their ability to complete presentations without AI, while simultaneously evaluating GenAI tools less positively. They may therefore be interpreted as dependent but less confident users. Because they appear to rely on the technology without being fully convinced of its value or ease of use, this group underscores the necessity of targeted pedagogical interventions. Educators must provide these students with structured practice to build both their independent presentation skills and their digital confidence.

This pattern is consistent with broader concerns in the literature that some students may rely on AI without fully trusting or valuing it. Their needs are unlikely to be addressed by either general pro-AI messages or strict usage restrictions, since they are not highly enthusiastic but also feel unable to avoid AI. Future longitudinal research could examine whether students' GenAI engagement develops from limited use toward broader and more confident integration. This has important implications for curriculum design, as early AI literacy support may influence the kind of GenAI engagement pathway students eventually develop.

5.3. Students' Evaluation of GenAI Usefulness and Ease of Use

The findings show that students evaluated GenAI positively in terms of both perceived usefulness and perceived ease of use. Both construct means were above the scale midpoint, suggesting that students generally found GenAI helpful and manageable for academic presentation preparation. However, the item-level results show an important distinction. Students rated GenAI most highly for helping them complete presentations faster and improve content or ideas. In contrast, they gave a lower rating to the item about whether AI clearly

improved their own presentation skills. This suggests that students valued GenAI more for immediate productivity than for long-term skill development.

The study's findings regarding students' evaluations of Perceived Usefulness can be further understood through the extended Technology Acceptance Model (TAM2) proposed by Venkatesh and Davis (2000). According to their theoretical framework, users form usefulness judgments through cognitive instrumental processes, consciously comparing what a technology is capable of doing with what they specifically need to achieve. In this study, students appeared to value GenAI mainly for practical support, especially speed and content improvement. However, their lower rating for presentation-skill development suggests that they may not view AI assistance as equally valuable for deeper learning growth. This distinction indicates that GenAI may support task completion more readily than skill development unless educators design activities that require reflection, revision, and critical evaluation of AI-generated output.

5.4. Pedagogical Implications

The findings suggest several implications for GenAI integration in academic presentation tasks. First, because students already use GenAI frequently, educators should provide clearer guidance on appropriate task delegation, output evaluation, and disclosure of AI assistance. Second, assessment design should recognize students' tendency to use AI more for preparation than for live performance by combining AI-supported preparation tasks with authentic oral delivery, explanation, and Q&A. Third, the three user profiles indicate that students may need differentiated support. Power Integrators may benefit from advanced guidance on critical evaluation and responsible multimodal use, while Selective Minimalists may need support in expanding effective uses of AI without weakening autonomy. Skeptical Delegators may require more structured practice to build confidence and reduce unreflective dependence. Finally, the gap between productivity gains and perceived skill development suggests that AI-assisted assignments should include reflection and revision activities so that students can connect AI use with learning growth rather than task completion alone.

5.5. Limitations

Several limitations should be considered when interpreting the findings. First, the study used a cross-sectional design, so it cannot establish causal relationships among AI use, perceived usefulness, ease of use, and intention to continue use. Second, the data were based on self-reported survey responses, which may not fully reflect students' actual AI practices, prompt behavior, or presentation performance. Third, the sample was drawn from one Vietnamese university, limiting the generalizability of the findings to other institutional, disciplinary, or cultural contexts. Fourth, the Behavioral Intention construct was measured with only two items, which may limit its conceptual coverage despite acceptable reliability. Finally, the K-means clustering results should be interpreted as exploratory, as cluster solutions can be sensitive to sample characteristics and require validation with independent datasets. Future studies should use longitudinal, mixed-method, or performance-based designs to examine how GenAI use affects presentation development over time.

6. Conclusion

This study examined undergraduate students' self-reported use of GenAI for academic presentation design and their evaluations of its usefulness and ease of use. The findings show that GenAI has become a routine support tool in students' presentation preparation, particularly for content-related tasks. However, frequent use did not necessarily indicate full dependence, as most students reported only moderate levels of task delegation and retained greater control over performance-related activities.

The study also shows that students are not a homogeneous group of GenAI users. The three profiles, Power Integrators, Selective Minimalists, and Skeptical Delegators, suggest that students differ in usage frequency, task delegation, tool diversity, perceived autonomy, and attitudes toward GenAI. In addition, students valued GenAI more strongly for immediate productivity and content improvement than for long-term presentation skill development.

These findings highlight the need for pedagogically guided GenAI integration in higher education. Rather than focusing only on whether students use AI, educators should support students in using it transparently, critically, and purposefully. AI-assisted presentation tasks should therefore include clear guidance on disclosure, appropriate delegation, output evaluation, and reflective revision, so that GenAI use contributes not only to task completion but also to students' academic and communicative development.

References

Baker, C., Ghassemi, E., & Bowers, R. (2024). Student perceptions of generative artificial intelligence in didactic patient presentations. *Pharmacy Education*, *24*(1), 590–597.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*, Article 43.

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments*, *10*, Article 60.

Chu, H.-C., Lu, Y.-C., & Tu, Y.-F. (2025). How GenAI-supported multi-modal presentations benefit students with different motivation levels: Evidence from digital storytelling performance, critical thinking awareness, and learning attitude. *Educational Technology & Society*, *28*(1), 250–269.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642.

Georgiev, S. G., & Tinsley, J. (2024). Exploring student acceptance and perceptions of AI-assisted PowerPoint creation. *African Journal of Inter/Multidisciplinary Studies*, *6*(S3), 1–13.

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, *30*, 1265–1300.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740–755.

Li, F., Abdullah, M. E. Z., & Tan, T. G. (2026). Exploring cognitive and motivational influences on students' acceptance of Artificial Intelligence Generated Content (AIGC) technology in product design instruction. *International Journal of Technology and Design Education*.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790.

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, *3*, 100076. https://doi.org/10.1016/j.caeai.2022.100076

Ngo, T. T. A. (2023). The perception by university students of the use of ChatGPT in education. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(17), 4–19.

Perkins, M. (2023). Academic integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, *20*(2), Article 07.

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, *32*(9), 5142–5155.

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, *6*(1), 1–10.

UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Bionote

Nguyễn Thị Hồng Liên is a university lecturer with expertise in TESOL, instructional design, and technology-enhanced learning. Her work focuses on English language education, AI-supported teaching, Moodle-based assessment, curriculum development, and learner-centered classroom practice. She is especially interested in helping teachers meaningfully integrate digital tools into teaching, assessment, and course design.

Nguyễn Thị Thanh Vân is an English lecturer at Dong Nai University. Her research interests include educational technology, teacher professional development, and English language teaching and learning.

#58 ASSESSING THE EFFECTIVENESS OF QUIZIZZ IN IMPROVING VOCABULARY LEARNING OF EFL FRESHMEN STUDENTS AT HUFLIT

LÊ THY PHƯƠNG THẢO

Abstract

Gamified learning instruments have received considerable focus in language teaching, especially in vocabulary development. This research examines the effectiveness of Quizizz in improving vocabulary acquisition among freshmen EFL learners at HUFLIT. Employing a combination of methods, information was gathered from 100 students via surveys and evaluations to assess the effect of Quizizz on students' enthusiasm, participation, and vocabulary retention. The results show a significant link between the use of Quizizz and increased student excitement, engagement, and academic outcomes. Nonetheless, constraints like monotonous question structures and the absence of customized learning pathways were noted. The research indicates that combining multimedia components, various question formats, and personalized learning capabilities could improve the platform's efficiency. These results underscore the significance of learning powered by technology strategies and offer perspectives for learners looking to enhance digital resources for vocabulary.

***Keywords*:** Gamified learning, Quizizz, vocabulary acquisition, EFL freshmen students, motivation, technology-enhanced learning, digital tools in education.

1. Introduction

1.1 Background

English vocabulary is essential for communication and professional success in today's globalized society. The need for creative and engaging ways to improve vocabulary learning continues to grow as English proficiency becomes increasingly crucial. Freshmen HUFLIT students who are prepared for their future occupations frequently struggle to stay motivated and successfully retain language using conventional approaches. This emphasizes the need for more effective and interactive educational resources. Quizizz, a gamified learning platform used in education to make learning more engaging and enjoyable, has seen increased popularity in recent years. However, its precise effect on enhancing freshmen HUFLIT students' vocabulary acquisition, particularly in terms of preparing them for professional contexts, remains an important topic of research. To fulfill the growing need for efficient vocabulary learning solutions for these students, this study looks at how Quizizz handles these difficulties.

1.2 Literature Review

To improve learning, English language teaching has undergone substantial change, utilizing cutting-edge resources and techniques (Harmer, 2007). These include e-learning programs like Quizizz, which have become well-known by allowing teachers to design popular online tests (Degirmenci, 2021, & Hidayati, 2019). A dynamic learning environment is promoted by Quizizz's game-like features, which encourage rivalry and engagement among students (Pagnattaro, 2017). Furthermore, Quizizz has been demonstrated to boost vocabulary acquisition through games, repetition, and images that enhance comprehension, pronunciation, and memorization (Nguyen Thi Huong, 2024). According to Hussain et al. (2023), as mentioned in Law et al. (2021), educational tools like Quizizz, Schoology, and Kahoot are commended for increasing motivation, interaction, prompt feedback, and overall learning efficiency. Additionally, interest is a vital learning catalyst that encourages people to commit

time and energy to their academic endeavors (Mesterjon et at., 2024, as referenced in Ratnasari, 2021).

1.3 Objectives

The purpose of the study is to determine how well Quizizz helps freshmen HUFLIT students acquire vocabulary and to make recommendations for methods to improve their motivation, retention, and preparedness for professional use. The results of a study carried out at HUFLIT to ascertain how Quizizz affects vocabulary acquisition and its function in preparing students for future occupations are presented in this research. This study's objective was to examine the effects of Quizizz on students' vocabulary acquisition, emphasizing quantifiable gains in their interest and performance. This thesis investigates whether Quizizz can help HULFIT freshmen students acquire the specific vocabulary required for success in both the classroom and the workplace.

1.4 Research Question

To ensure that the research corresponds to the aim, the study is guided by the following research question: In what way does Quizizz contribute to the vocabulary learning of EFL freshmen students of HUFLIT?

1.5 Scope

The present paper is devoted to the use of Quizizz by HUFLIT freshmen students for vocabulary learning. It evaluates the degree to which Quizizz helps to engage students and improve their vocabulary acquisition results. The study involves 100 freshmen EFL students who have been using Quizizz in vocabulary-learning activities before. The research takes place during March and April 2026, thus providing a time frame for data gathering and analysis. The possible biases in self-reported data and the differences in participants' previous experiences with gamification platforms such as Quizizz are some limitations of the study.

1.6 Organization

The present paper is structured into five main sections. The first section includes an introduction, aim, research questions, scope, and organization of the study. The second one reviews the literature related to vocabulary learning, gamified learning, and the application of Quizizz in the EFL context. The third section provides information about research design, participants, data-gathering tool, procedure, and statistical treatment of data. The fourth section states the results concerning the effectiveness of Quizizz, students' engagement, helpful features, and challenges. The fifth section contains a discussion of results, limitations, implications, and recommendations.

2. Literature review

2.1 Vocabulary Learning

Since vocabulary is the basis for both successful communication and the development of other language abilities, it is essential for language learning. To improve vocabulary acquisition, Harmer (2007, p.229) highlights that students should study words in context and comprehend both their meaning and structure. Improving vocabulary development is still a major goal in English language programs, especially at the fundamental level, because of its significance. Understanding spoken and written language requires a strong vocabulary (Boyinbode, 2018). Learners find it difficult to understand and effectively communicate concepts when they are unaware of word definitions.

According to Afzal (2019), vocabulary is essential for language acquisition and communication. He points out that students' vocabulary acquisition and academic performance can be greatly improved by employing cognitive and metacognitive techniques in addition to providing learning opportunities. Ghalebi et al. (2021) stated that vocabulary development is crucial for learning a foreign language since vocabulary difficulties can impede communication and reading comprehension, especially for non-native English speakers.

Yawiloeng (2020) emphasizes once more how important vocabulary acquisition is to the development of fundamental language abilities, such as speaking, listening, reading, and writing. Teachers must use efficient teaching techniques to get students interested in acquiring language. According to Teng (2019), the objective of vocabulary learning is for students to commit new words to memory and use them in a variety of conversational scenarios. However, because vocabulary must be actively used in productive abilities like speaking and writing, some students find this process tiresome because of the demands of memory and repetition.

Early exposure to new words greatly improves language development; therefore, young learners must master vocabulary (Ramadhan & Zaharani, 2021). Vocabulary acquisition should be given top priority to promote general language competency and fluency.

2.2 Significance of Vocabulary Learning in English

Before they may utilize new words in speech and improve their language skills, learners must be exposed to them repeatedly through listening (Rowe, 2012). Thus, it is essential to teach vocabulary in the early stages of language learning (Schmitt, 2010). To put it another way, to demonstrate their language skills, learners need to get familiar with a particular collection of vocabulary. It would be challenging to communicate thoughts and comprehend others without a large vocabulary.

By increasing their productive vocabulary, English language learners can develop their communication abilities and achieve fluency and clear expression in the target language. Kilic (2019) discovered that among English as a foreign language learners, vocabulary knowledge is a strong predictor of speaking and writing performance, underscoring the significance of increasing vocabulary to improve communication abilities. This implies that having a strong vocabulary not only makes comprehension easier, but it is also essential for students to be able to express themselves clearly both orally and in writing.

2.3 Challenges in Vocabulary Learning

Even though studying vocabulary is crucial to learning English as a foreign language (EFL), many students encounter a variety of challenges. These difficulties include learning new words, pronouncing, using, spelling, and memorizing them correctly (Afzal, 2019). Additionally, acquiring synonyms, distinguishing words with numerous meanings, and appropriately utilizing prefixes and suffixes memorizing them (Afzal, 2019, p. 81). Additionally, acquiring synonyms, distinguishing words with numerous meanings, and appropriately utilizing prefixes and suffixes are all challenges that EFL students frequently face (Khan, 2011, p. 1255).

Errors among university students have also been noted in other research (Abdulaal, 2019; Sheir et al., 2016), especially confusion brought on by near-synonyms and similar spellings. These results show that EFL learners continue to struggle with vocabulary acquisition, which impedes their overall language ability. Given the significance of language acquisition, academics have looked for novel ways to improve instruction beyond conventional techniques. Vocabulary education is one of the new instruments for English language teaching

brought about by technological advancements. Multimedia software can be included in language instruction to enhance vocabulary acquisition (Astika, 2014).

Recent advancements in technology offer teachers and students unmatched opportunities for vocabulary acquisition (2017, p.45). By boosting student involvement and offering relevant settings for word comprehension, digital tools and media can improve vocabulary learning (Hermagustiana & Resmawaty, 2017, p. 138). In response to these issues, this study looks at how educational apps like Quizizz can help EFL students acquire vocabulary more effectively.

2.4 Quizizz's Effects on Freshmen's Motivation and Participation in Vocabulary Learning

The millennial generation, which is strongly associated with technology and contemporary developments, is particularly affected by this problem. Gamification has become a successful tactic in language classes, and integrating technology into language learning is essential for closing this gap (Medina & Hurtado, 2017).

As a digital learning tool, gamification is flexible and provides individualized feedback based on the requirements of every student. It facilitates differentiated instruction and aids in students' vocabulary acquisition through repetition and context (Abrams & Walsh, 2014). Quizizz is one such gamified tool that enables educators to design captivating, creative, and interactive lessons using kinesthetic and audiovisual components (Aini, 2019).

Teachers can increase student motivation and engagement by integrating Quizizz into vocabulary instruction, which will make learning more pleasurable and productive. Quizizz is an effective teaching tool that uses gamification to improve vocabulary acquisition in English. Its dynamic and captivating features are intended to inspire students and encourage their engagement with the material covered in class.

In contrast to conventional formative assessments, Quizizz offers a fun and competitive setting that discourages cheating while enhancing the learning experience (Amalia, 2020). Additionally, it helps teachers evaluate students' language skills and curriculum comprehension (Bury, 2017).

Quizizz's game-based features, which include avatars, themes, memes, and music, are one of its main characteristics that make studying entertaining and interesting. By reviewing vocabulary interactively, students can increase their engagement and recall (Pavita & Nirmala, 2021). Quizizz is a fun way to kill time. Students may learn at any time and from any location thanks to the platform's accessibility on a variety of platforms, including PCs, laptops, tablets, and smartphones.

Students can find errors and strengthen their vocabulary by using the platform's instant feedback (Nhu & Tin, 2019). According to research, students feel favorably about utilizing Quizizz as an online assessment tool, especially for formative tests (Pusparani, 2021). Quizizz also provides teachers with comprehensive information, such as reports at the class and student levels that may be saved as Excel files for additional analysis.

2.5 Advantages of Using Quizizz

Quizizz has several advantages for teaching and studying English, especially when it comes to grammatical comprehension, motivation, and self-evaluation. Rahayu and Purnawarman (2019) claim that Quizizz helps students evaluate their strengths and weaknesses, thereby enhancing their academic performance through frequent tests,

evaluations, and feedback. According to Razali et al. (2020), Quizizz's gamification features are essential for encouraging students' internal and extrinsic motivations.

Bal (2018) found that Quizizz improves vocabulary retention more effectively than conventional approaches. Similarly, Quizizz greatly enhances grammar learning, increasing the effectiveness of knowledge acquisition, according to Dewi et al. (2020). Additionally, according to Lestari's (2019) comparison of Quizizz and Kahoot, students prefer Quizizz because of its better levels of engagement, challenge, and incentive. Additionally, Quizizz encourages students to adopt a more proactive approach to learning by creating a positive washback effect, according to Pitoyo et al. (2020).

Overall, research results confirm that Quizizz is not only a useful instrument for teaching English but also a strong motivator that improves students' vocabulary and grammatical acquisition while fostering an interesting and fruitful learning environment.

2.6 Disadvantages of Using Quizizz

Quizizz has drawbacks despite its many benefits. One disadvantage is that students' scores may be impacted by how quickly they complete the test; the quicker they respond, the higher their scores, which might not necessarily accurately represent their comprehension. Furthermore, Quizizz is dependent on a reliable internet connection; if the connection is poor and unreliable, students may be locked out, which can lead to lost progress.

The fact that some advanced functions, including embedding instructional films, need a premium subscription is a major drawback for educators. This might prevent the full potential of the free version. Internet dependence is a significant obstacle to adopting Quizizz as an assessment tool.

Education has been greatly impacted by the growing availability of the Internet in homes, companies, and educational institutions. Research shows that students' ability to access online material has a significant impact on their academic achievement. Smartphones are becoming the most common way to access the Internet, which facilitates easy contact among academics and supports learning, research, and assignment completion.

3. Methodology

3.1 Overview

This study examines how well Quizizz helps freshmen EFL students at HUFLIT learn vocabulary and how it relates to their future employment. To fully comprehend students' learning experiences, a mixed-methods approach was used, combining quantitative and qualitative data. A questionnaire that evaluated vocabulary learning difficulties, students' involvement with Quizizz, and their opinions of its efficacy served as the main instrument for gathering data.

3.2 Sampling Technique

Convenience sampling was used since it is a realistic way to effectively reach a large number of individuals. One hundred HUFLIT freshmen EFL students were chosen to take part in this study. Students who actively participated in Quizizz vocabulary-learning activities were chosen as participants. Although the results' generalizability was restricted by the convenience sampling technique, it offered insightful information on the experiences of a particular set of students.

3.3 Sample Restriction

To ensure consistency in language proficiency, the study specifically targeted freshmen EFL students at the intermediate level. Students who were not actively involved in vocabulary learning with Quizizz were excluded from the study to maintain reliability in data collection and analysis. By focusing on this specific group, the study aimed to provide accurate and meaningful insights into the effectiveness of Quizizz in vocabulary acquisition.

3.4 Materials

A structured questionnaire built with Google Forms served as the study's main method of data collection. Each of the four sections of the questionnaire was intended to cover a different facet of the study. Name, age, and gender were among the demographic information gathered in the background information section. Multiple-choice questions were used in the second part. The third section, Engagement with Quizizz, utilized Likert-scale items to evaluate students' motivation and interaction levels. The fourth section, Perceived Effectiveness of Quizizz, featured rating scales and open-ended questions to gather insights into students' experiences and opinions on how Quizizz contributed to their vocabulary learning. Before distribution, the questionnaire underwent a review process to ensure clarity, reliability, and relevance.

3.5 Procedure

The research was conducted systematically in several stages to ensure a thorough investigation into the impact of Quizizz on vocabulary acquisition. The study was carried out from March to April of 2026, during which data collection took place. The first stage involved questionnaire design and validation. The research team carefully constructed the survey, ensuring that the questions aligned with the study's objectives. Once completed, the questionnaire was reviewed by language educators and peers to confirm its clarity and relevance. In the second stage, the distribution of the questionnaire, the finalized survey was shared with participants via a private online group of students. Participants were given detailed instructions on how to complete the questionnaire, ensuring that they understood each question type, including multiple-choice questions, Likert-scale items, rating scales, and open-ended responses. Students were allocated 5-6 minutes to complete the survey at their convenience. The third stage focused on data collection and monitoring. The research team monitored responses to ensure a sufficient number of completed questionnaires. The participation rate was observed daily, and reminders were posted in the online group to encourage full engagement. After reaching the target number of responses, the final stage involved data compilation and preparation for analysis. In April 2026, the questionnaire was closed, preventing further submissions. The collected responses were then organized, cleaned, and prepared for statistical and thematic analysis. By following this structured procedure, the study ensured the reliability and validity of the data collected.

3.6 Statistical Treatment

A combination of quantitative and qualitative analysis methods was applied to analyze the collected data. Quantitative responses, including multiple-choice and Likert-scale items, were analyzed using descriptive statistical techniques such as frequency distribution, mean scores, and percentage analysis. This approach provided numerical insights into students' engagement levels,

learning experiences, and perceived effectiveness of Quizizz. Open-ended responses were analyzed qualitatively to identify patterns, themes, and subjective insights regarding students' perceptions of Quizizz. The mixed-methods approach ensured a well-rounded evaluation of the impact of Quizizz on vocabulary acquisition, integrating numerical data with qualitative feedback for deeper analysis and interpretation.

4. Results

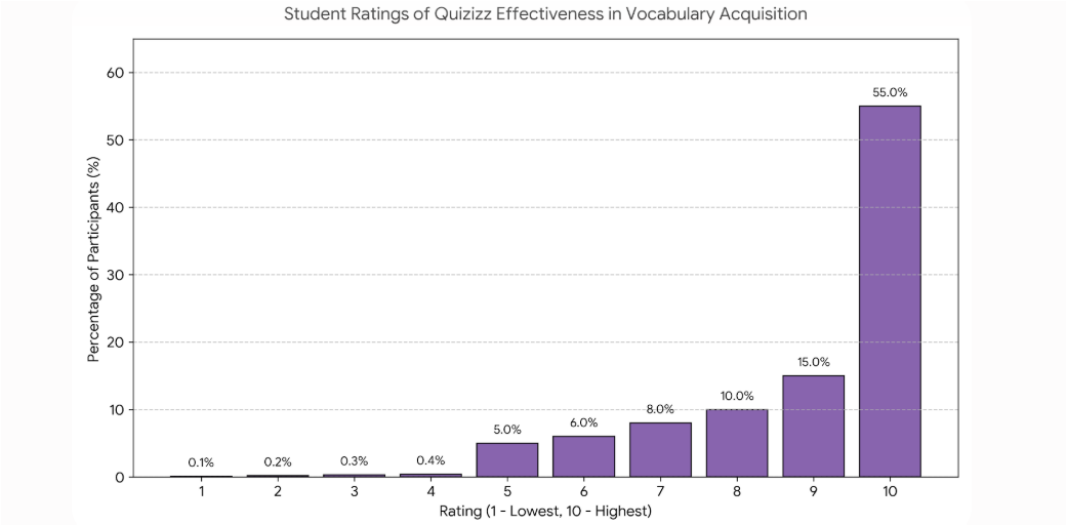
4.1 Overview of Findings

An overview of the respondents' demographics, including gender distribution and educational background, is given in this section. Knowing these demographics guarantees a variety of viewpoints when evaluating Quizizz's efficacy and aids in placing the results in context.

4.2 Effectiveness of Quizizz in Vocabulary Learning

Figure 1

Student Ratings of Quizizz Effectiveness in Vocabulary Acquisition

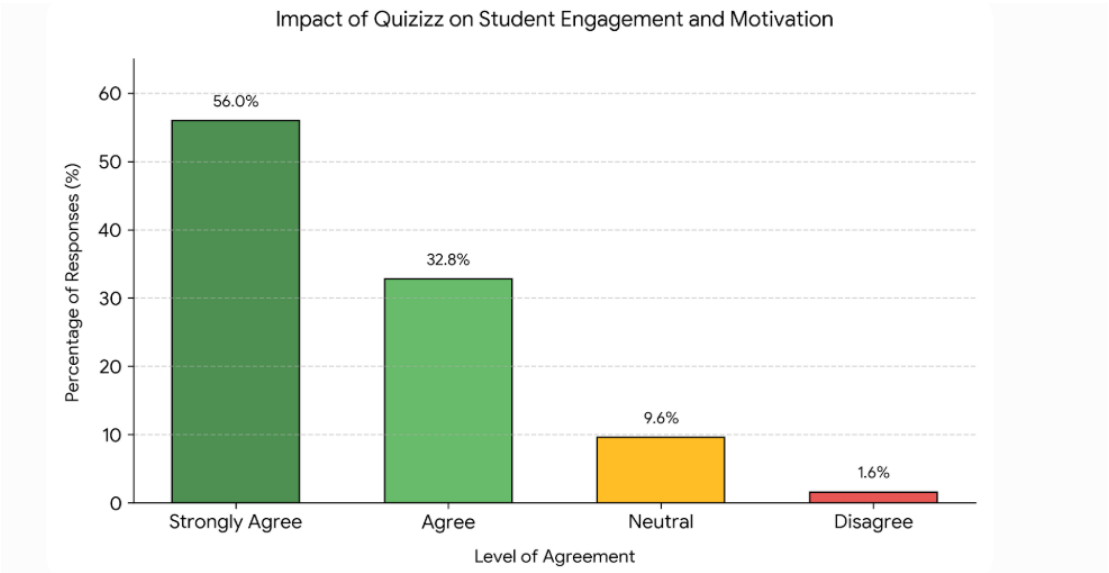


More than half of the participants gave the platform the highest possible rating of ten, while about a quarter of the group rated it slightly lower at nine or eight. A small percentage of students viewed its effectiveness as moderate, scoring it at a seven or six, and roughly one in twenty students gave it a neutral rating of five. On the other hand, dissatisfaction was remarkably low, with fewer than one percent of students assigning a rating of four, three, or two. Overall, these strongly positive reactions highlight that Quizizz successfully supports and enhances vocabulary acquisition for the clear majority of learners.

4.3 Engagement and Enjoyment

Figure 2

Impact of Quizizz on Student Engagement and Motivation

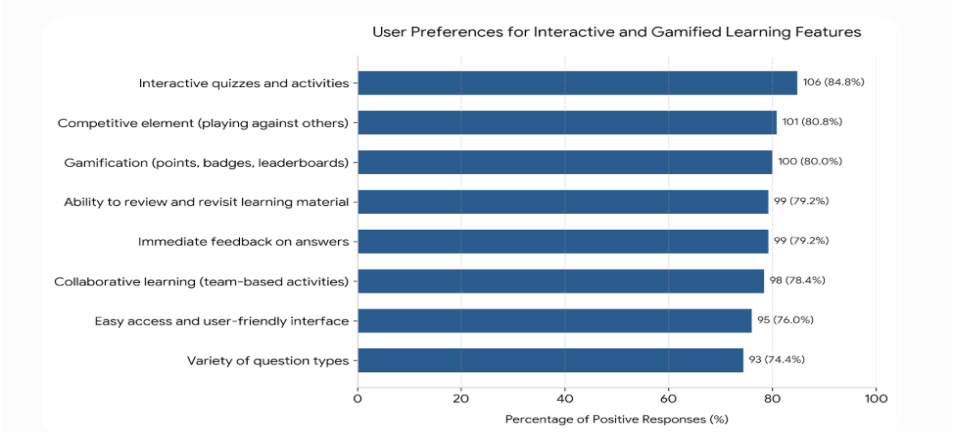


According to the findings, Quizizz greatly increased engagement. There was an 88.8% positive response rate, with 56% of students strongly agreeing and 32.8% agreeing that Quizizz made studying vocabulary more interesting and fun. In the meantime, a small portion disagreed, and 9.6% stayed neutral. This demonstrates how the platform's gamified and interactive features are essential for boosting motivation and making vocabulary learning more pleasurable.

4.4 Features that Enhance Vocabulary Learning

Figure 3

User Preferences for Interactive and Gamified Learning Features



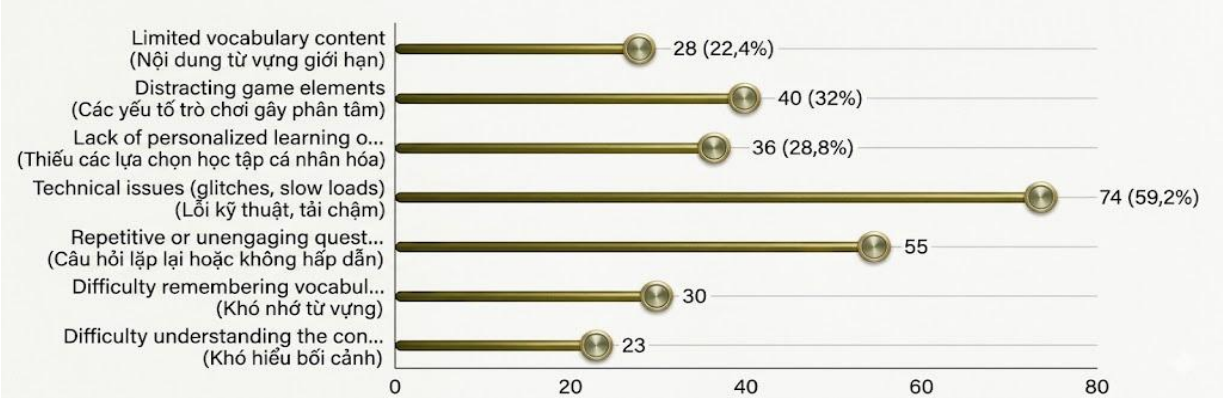
The horizontal bar chart, titled "User Preferences for Interactive and Gamified Learning Features," presents data on what users find most valuable in a modern learning environment, measured by both total count and percentage of positive responses. Out of the surveyed features, "Interactive quizzes and activities" ranks as the most preferred element, securing the highest approval at 106 responses (84.8%). Strong engagement is also seen in competitive and reward-based features, with "Competitive element (playing against others)" closely following at 101 responses (80.8%) and "Gamification (points, badges, leaderboards)" at 100 responses (80.0%). Review and feedback mechanisms also show high, identical utility among learners; both "Ability to review and revisit learning material" and "Immediate feedback on answers" tied at 99 responses (79.2%), while "Collaborative learning (team-based activities)" sits just a

fraction lower at 98 responses (78.4%). Rounding out the bottom of the preferences are "Easy access and user-friendly interface" with 95 responses (76.0%) and "Variety of question types" with 93 responses (74.4%). Overall, the chart demonstrates a consistently strong user demand for all listed interactive tools, with even the lowest-ranked feature enjoying an overwhelming popularity rate well above 70%.

4.5 Challenges in Using Quizizz

Figure 4

Challenges in Using Quizizz

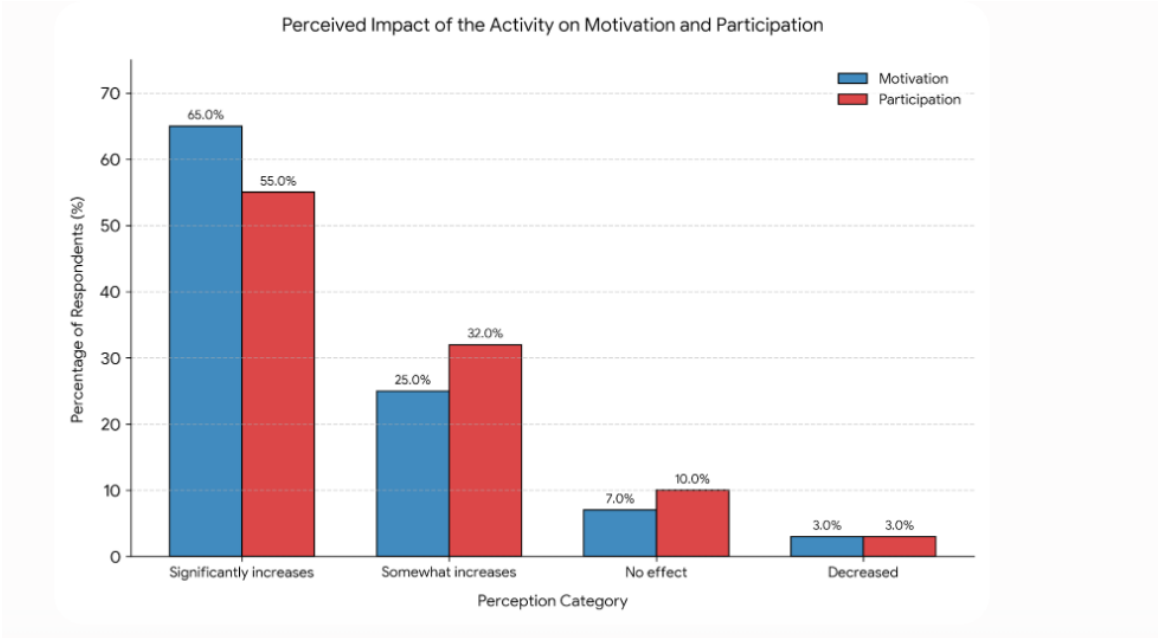


The horizontal bar chart details various challenges and drawbacks encountered by users, detailing both the frequency and percentage of responses for each category. "Technical issues (glitches, slow loads)" stands out as the most prominent problem by a significant margin, reported by 74 respondents (59.2%). This is followed by concerns regarding content and gameplay design, where "Repetitive or unengaging questions" affected 55 users, and "Distracting game elements" were cited by 40 participants (32%). In terms of structural limitations, a "Lack of personalized learning options" was flagged by 36 respondents (28.8%), while "Limited vocabulary content" was noted by 28 users (22.4%). Finally, user-specific learning difficulties were less frequent but still present: 30 participants reported "Difficulty remembering vocabulary", and 23 users noted "Difficulty understanding the context" (18.4%). Collectively, the data indicate that software reliability and content engagement are the primary hurdles users face, while individual comprehension difficulties remain minor pain points by comparison.

4.6 Respondents' Perceptions of How the Activity Affects Their Motivation and Participation.

Figure 5

Perceived Impact of the Activity on Motivation and Participation



The chart illustrates respondents' perceptions of how the activity affects their motivation and participation. A large majority of participants reported that the activity significantly increases both their motivation and participation, with the highest percentages shown in the blue bars. Additionally, a considerable number of respondents indicated that the activity somewhat increases these aspects, as represented by the red bars. Only a small proportion believed that the activity had no effect, while very few respondents felt it decreased their motivation or participation. Overall, the findings suggest that the activity has a strongly positive impact on both motivation and participation.

5. Conclusion and discussion

5.1 Research Question Revisited

This study addressed the main research question by evaluating the extent to which Quizizz supports vocabulary learning among freshmen EFL students at HUFLIT. It also examined how Quizizz affects students' motivation and participation, which features students find most useful, and what challenges they encounter when using the platform for vocabulary learning.

5.2 Findings

The findings show that Quizizz's interactive and gamified features greatly improve student engagement and vocabulary retention. The study shows that using Quizizz increases student motivation, with many participants expressing a stronger desire to acquire language. The site also offers instant feedback, which promotes active recall and self-evaluation. However, some students encountered challenges such as repetitive question formats and limited opportunities for personalized learning. Despite these drawbacks, the findings confirm that Quizizz serves as an effective tool for vocabulary acquisition.

5.3 Explanation of Findings

The findings are consistent with earlier research highlighting the importance of gamification in language acquisition. Quizizz's competitive features, which boost motivation and strengthen vocabulary recall through repeated exposure, are responsible for its engagement-driven character. Additionally, by enabling students to track their development

and correct errors instantly, the quick feedback system promotes metacognitive techniques. However, some challenges, such as technical issues and repetitive content, may hinder the learning experience. These limitations suggest that while Quizizz is beneficial, its effectiveness can be optimized by incorporating more diverse question formats and adaptive learning features.

5.4 Limitations

The generalizability of the results is limited by the sample size, as the study only included freshmen EFL students at HUFLIT. Additionally, external factors such as students' prior experience with gamified learning and individual learning preferences might have influenced the findings. The study also did not account for long-term vocabulary retention beyond the immediate testing period. Future research should consider a broader sample size and a longitudinal approach to assess the sustained impact of Quizizz on vocabulary development.

5.5 Discussion

5.5.1 Implications

The results highlight the importance of using technology-driven learning resources in language instruction. Students' favorable reactions to Quizizz demonstrate how gamified platforms can improve learning results and engagement. Additionally, the study contends that interactive learning promotes self-directed learning habits and autonomy, both of which are critical for long-term language acquisition. To get the most out of Quizizz, teachers need to incorporate extra activities like group discussions and writing assignments to guarantee greater vocabulary understanding and real-world applications.

5.5.2 Recommendations

To optimize the effectiveness of Quizizz in vocabulary learning, educators should integrate it into formal instruction by aligning quizzes with course objectives. Enhancing Quizizz with multimedia elements such as images, audio pronunciations, and contextual examples can improve vocabulary retention and pronunciation accuracy. Diversifying question formats to include fill-in-the-blank, sentence construction, and open-ended responses would encourage critical thinking and deeper understanding. Additionally, incorporating personalized learning paths within Quizizz can cater to individual student needs, allowing for targeted vocabulary reinforcement. Finally, conducting follow-up assessments to measure vocabulary retention over time would provide valuable insights into the platform's long-term impact. By implementing these recommendations, Quizizz can become a more dynamic and effective tool for vocabulary learning, equipping students with essential linguistic skills for academic and professional success.

References

Afzal, N. (2019). A study on vocabulary-learning problems encountered by BA English majors at the university level of education. *Arab World English Journal*, 10(3), 81– 98. https://doi.org/10.24093/awej/vol10no3.6

Alhebshi, A. A., & Gamlo, N. (2022). The effects of mobile game-based learning on Saudi EFL foundation year students' vocabulary acquisition. *Arab World English Journal, 13*(1), 408–425. https://doi.org/10.24093/awej/vol13no1.27

Amalia, D. F. (2020). Quizizz website as an online assessment for English teaching and learning: Students' perspectives. *Journal of English Language Teaching*, 7(1), 1-8. http://ojs.ikipmataram.ac.id/index.php/joelt

Ameen, F. (2026). Individual and group competitive digital gamification in ESL adult classrooms: A systematic review of effects on vocabulary retention, motivation, and engagement. *Smart Learning Environments, 13*, Article 22. https://doi.org/10.1186/s40561-026-00447-z

Handoko, W., Mizkat, E., Nasution, A., Hambali, & Eska, J. (2021). Gamification in learning using Quizizz application as assessment tools. *Journal of Physics: Conference Series, 1783*(1), 012111. https://doi.org/10.1088/1742-6596/1783/1/012111Harmer, J. (2007). *The practice of English language teaching.* Harlow: *Longman.*

Hsu, H.-L. (2025). *The effect of combined Wordwall games' types on EFL learners' vocabulary learning and learner engagement.* RELC Journal. *https://doi.org/10.1177/00336882251396580*

Jia, W., Zhang, L., Pack, A., Guan, Y., & Zou, B. (2024). *Digital game-based learning's effectiveness on EFL learners' receptive and productive vocabulary knowledge. Language Learning & Technology, 28(1), 1–21. https://doi.org/10.64152/10125/73554*

Khan, I. A. (2011). Learning difficulties in English: Diagnosis and Pedagogy in Saudi Arabia. Educational Research, 2(7), 1248-1257. https://www.researchgate.net/publication/234168115

Li, R. (2021). Does game-based vocabulary learning APP influence Chinese EFL learners' vocabulary achievement, motivation, and self-confidence? *SAGE Open*. https://doi.org/10.1177/21582440211003092

Pusparani, A. D. (2021). The Quizizz platform is the realization of asynchronous learning to improve students' vocabulary mastery. ELT Echo: The Journal of English Language Teaching in Foreign Language Context, 6(2). https://doi.org/10.24235/eltecho.v6i2.9312

Razali, N., Nasir, N. A., Ismail, M. E., Sari, N. M., & Salleh, K. M. (2020). Gamification elements in Quizizz applications: Evaluating the impact on intrinsic and extrinsic student motivation. *IOP Conference Series: Materials Science and Engineering, 917*(1), 012024. https://doi.org/10.1088/1757-899X/917/1/012024

Wu, T.-T., & Huang, Y.-M. (2017). A mobile game-based English vocabulary practice system based on portfolio analysis. *Educational Technology & Society, 20*(2), 265–277.

#59 Self- Reflection in Using GenAI for Learning: Perceptions of English-Majored Students at Binh Duong University- Ca Mau Campus

Liet Thoai Phuong Lan [1]

Abstract

Generative AI (GenAI) has revolutionised higher education, especially in English Language Studies. However, GenAI is most effective when it supports students in developing their own thinking. Therefore, this study examines the perceptions and self-reflection practices of English-major students regarding the use of GenAI in their academic paths. There were four keys: academic problem-solving, creative expression, communication, and decision-making, which were used to structure the self-reflection checklist. The research investigates how students balance the use of GenAI assistance with independent thinking across these key areas. The author employed a mixed-methods approach, using questionnaires for quantitative data collection and reflective journals and semi-structured interviews for qualitative data collection. Data were collected from 50 English major students at Binh Duong University - Ca Mau Campus. The results indicate that GenAI is widely used for drafting and refining ideas, but there is a clear need to develop strategic autonomy to avoid creative stagnation and over-dependence. Moreover, it highlights the importance of self-reflection as a pedagogical tool to promote ethical and effective AI integration in English Language Studies.

***Keywords*:** GenAI, self-reflection, English majors, learner autonomy, academic integrity

1. Introduction

Today, Artificial Intelligence (AI) has become useful in higher education. Generative Artificial Intelligence (GenAI) tools such as ChatGPT, Gemini, and Claude are increasingly used in educational settings, raising questions about their impact on student learning (Xiaodong Chen et al., 2026). Additionally, Kasneci et al. (2023) show that tools like ChatGPT can assist with language production, provide immediate feedback, and encourage learner independence. Teaching and learning have changed significantly with the adoption of GenAI. Nina Rizun et al. (2026) describe that ChatGPT and similar platforms are often used for idea generation, writing support, and language improvement in English Language Studies. Owing to these technological advantages, students are relying more on AI, which could hinder the development of critical thinking skills (Sofie Otto et al., 2025; Nina Rizun et al., 2026). These researchers warn that GenAI might create a false sense of ability, where students produce decent work without truly understanding the underlying processes. This trend raises important questions about student engagement and responsibility in AI-supported learning environments. To address these concerns, self-reflection has become an essential tool for meaningful learning. Instead of viewing GenAI as a substitute for critical thinking, self-reflection encourages students to evaluate when, why, and how they use AI tools. Therefore, this study contributes to the literature in three ways. Initially, it examines self-reflection as a mediating mechanism that shapes whether GenAI promotes learner autonomy or fosters dependency. In addition, it introduces a structured self-reflection framework that captures students' behaviours across four domains: academic problem-solving, creative expression, communication, and decision-making. More critically, it provides evidence from English-

[1] *Master of English Linguistics, Faculty of Foreign Languages, Binh Duong University-Ca Mau campus, Ca Mau City, Vietnam; Email: ltplan.cm@bdu.edu.vn*

major students in Binh Duong University – Ca Mau campus, which remains underrepresented in current AI-in-education research. By focusing on reflective engagement rather than merely frequency of AI use, the study offers a more nuanced understanding of sustainable and ethical AI-supported learning.

2. Literature Review

2.1. Self-Reflection and Learner Autonomy

Self-reflection is essential to metacognition and plays a crucial role in encouraging learner independence in higher education. Research shows that self-regulated learning involves continuous processes of planning, monitoring, and reflecting, enabling students to oversee their cognitive and behavioural actions (Zimmerman, 2000; Zimmerman, 2002; Panadero, 2017). Reflection helps students assess their performance and adjust their strategies, thereby enhancing learning effectiveness (Schön, 1983; Boud, Keogh & Walker, 1985). For example, structured reflection boosts students' motivation, self-confidence, and performance. These findings suggest that self-reflection is vital for meaningful learning, especially when learners encounter complex cognitive tasks.

2.2. Generative AI in Language Learning: Integrating

Generative Artificial Intelligence (GenAI) in education has created new opportunities for personalised learning. Students receive immediate feedback, brainstorm ideas, and get help with writing tasks, making these tools especially useful in English Language Studies through AI-based language models. Kasneci et al. (2023) argue that AI tools can support tailored learning paths and boost student engagement. Similarly, Zawacki-Richter et al. (2019) offer a comprehensive review showing that, when used correctly, artificial intelligence can improve teaching and learning. However, concerns about over-reliance on AI tools have also emerged. Debby R. E. Cotton et al. (2024) note that students may rely on AI to complete tasks without fully understanding the content, leading to superficial learning. This issue is especially important because writing is both a mental and creative process in language learning. Relying too much on GenAI could prevent the development of authentic voice, critical thinking, and language skills.

2.3. Self-Reflection in the Use of GenAI

The effectiveness of GenAI in education depends not only on access to technology but also on how learners engage with it. Self-reflection is essential in ensuring that AI tools serve as cognitive support rather than replacements for thinking. Ling Wei (2023) provides valuable insights for educators and researchers interested in incorporating AI-powered platforms into language classes. The findings support the idea that AI-enhanced language instruction can transform learning and highlight the positive influence of AI-driven educational tools. Similarly, the authors (2023) emphasise the importance of AI literacy, including the ability to critically evaluate AI outputs and recognise their limitations. In this context, structured reflection tools—such as self-reflection checklists—can be effective teaching strategies. These tools assist students in assessing their behaviours, identifying over-reliance, and developing more strategic approaches to AI use. Through detailed thematic analysis, this study identified both the advantages and disadvantages of AI-driven language education, highlighting AI's role in fostering self-regulated learning while addressing challenges such as technological dependence and equitable access.

2.4. Conceptual Framework

Building on the existing literature on Self-Regulated Learning, this study introduces a conceptual framework to analyse how students interact with GenAI while learning English. In this framework, GenAI use is viewed as a double-edged sword that can either support or hinder learning depending on how it is mediated through self-reflection. While GenAI offers benefits like idea generation, language correction, and feedback, unregulated use may lead to passive learning and over-reliance. Additionally, self-reflection plays a crucial role in influencing the effectiveness of GenAI use. Through reflective practices, learners' interactions with AI tools adjust their learning strategies and take ownership of their cognitive processes. Without proper training, these activities align with the self-reflection stage of the Self-Regulated Learning model, where learners evaluate outcomes and implement necessary changes. Specifically, learner autonomy is defined as the main outcome variable, representing students' ability to think independently, maintain their unique voice, and take responsibility for their learning. At the same time, the study introduces the concept of creative dependency as a negative outcome, describing a situation where students struggle to generate ideas or complete tasks without AI assistance.

In summary, the proposed model indicates that GenAI use influences both learner autonomy and creative dependence, but these effects are strongly affected by the level of self-reflection. High self-reflection leads to a strategic and effective use of GenAI, whereas low self-reflection increases the likelihood of dependence on GenAI.

2.5. Research Gap

Despite the rapid growth of research on Generative Artificial Intelligence (GenAI) in education, several important gaps remain unaddressed, particularly in English Language Studies. To begin with, most studies focus on the benefits and risks of GenAI, such as increased productivity, personalised feedback, and concerns about academic integrity, without thoroughly exploring the learning processes behind students' interactions with these tools. As a result, the existing literature often evaluates outcomes—such as performance and efficiency—rather than the cognitive and metacognitive mechanisms that support meaningful learning. Subsequently, although self-regulated learning and learner autonomy are well-studied in language education, their connection with GenAI use remains underexplored. Specifically, the role of self-reflection as a mediator between GenAI use and learning outcomes remains underexplored. Most previous studies treat AI use as a direct predictor of learning results, overlooking how reflective practices may enhance or limit their effectiveness. Of particular importance is the finding that there are no valid, practical tools to measure students' strategic versus dependent use of GenAI. Current research rarely operationalises learners' behaviours across multiple domains—such as academic problem-solving, creative expression, communication, and decision-making—making it difficult to identify subtle patterns, such as "creative dependency." Ultimately, data from non-Western contexts, especially Vietnam, is limited. Given the sociocultural and educational differences that influence technology use and learner autonomy, findings from Western settings cannot be easily generalised. Therefore, research focused specifically on local contexts that reflect teaching practices and student characteristics is needed. Lastly, unlike previous studies that primarily investigated attitudes toward GenAI or its impact on learning outcomes, the present study focuses on the reflective processes that shape students' interactions with AI. By conceptualising self-reflection as a mediating mechanism rather than a simple learner characteristic, this study offers a process-oriented perspective on AI-supported learning. To address these gaps, the current study uses a mixed-methods approach. It introduces a structured self-reflection framework to examine how English majors engage with GenAI across different

areas. By viewing self-reflection as a mediating factor, the study aims to improve understanding of how GenAI can support or hinder learner autonomy and sustainable learning practices.

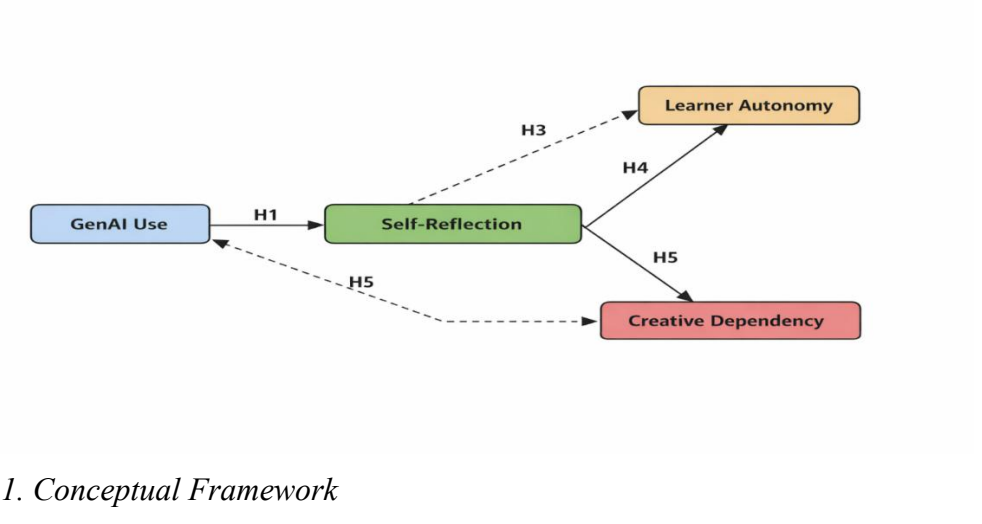


Image 2.1. Conceptual Framework

Note:

- *Solid lines (→) show the direct path of impact from behaviour (Gen AI use) → cognition (Self-reflection) →outcomes (Autonomy or dependency)*
- *Dashed lines (---->) illustrate the double-edged sword mechanism: self-reflection acts as a crucial mediator to achieve autonomy. At the same time, a lack of independent creativity can lead to an absolute reliance on AI.*

2.6. Research Hypotheses

According to the conceptual framework, the study suggests the following hypotheses:

H1: The use of GenAI is positively associated with students' perceived learning support among those majoring in English.

H2: Self-reflection is positively associated with learner autonomy.

H3: Self-reflection mediates the relationship between GenAI use and learner autonomy.

H4: Self-reflection has a negative relationship with creative reliance.

H5: Low levels of self-reflection are associated with increased use of GenAI in academic tasks.

3. Methodology

3.1. Research Design

This study used a mixed-methods research design, combining quantitative and qualitative approaches to gain a comprehensive understanding of students' GenAI use and reflection practices.

3.2. Participants

The participants included 50 English-major students at Binh Duong University – Ca Mau. Campus. These students had prior experience using GenAI tools in their academic work.

3.3. Instruments

(1) Questionnaire

A structured questionnaire was developed from a self-reflection checklist. It assessed students' behaviours across four areas: Academic problem-solving, Creative expression, communication, and Personal decision-making.

(2) Reflective Journals

Students were asked to record their experiences with GenAI, emphasising how they used the tool and how it affected their thinking.

(3) Semi-Structured Interviews

Interviews were carried out to gain a deeper understanding of students' perceptions and challenges.

3.4. Data Collection and Analysis

Quantitative data were analysed with descriptive statistics, while qualitative data were thematically coded. Combining both data types enabled triangulation and improved validity.

4. Results

4.1. Descriptive Statistics

Table 1

Descriptive Statistics of Key Variables

Variables	N	Mean	Std. Deviation
GenAI Use	50	4.12	0.68
Self-reflection	50	3.45	0.72
Learner Autonomy	50	3.58	0.65
Creative Dependency	50	3.21	0.74

Note. *Values based on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree).*

Table 1 presents the descriptive statistics of the key variables. The results show that students reported a high level of GenAI use (M = 4.12, SD = 0.68), indicating that AI tools are widely integrated into their academic routines. In contrast, self-reflection was at a moderate level (M = 3.45, SD = 0.72), suggesting that reflective engagement with GenAI is still developing. Additionally, learner autonomy was moderate (M = 3.58, SD = 0.65), while creative dependence was noticeable among students (M = 3.21, SD = 0.74). These findings point to a potential imbalance between technology use and independent thinking. Overall, the findings demonstrate that GenAI is extensively used by students, especially for idea generation, refining writing, and grammar checking. However, most students tend to use GenAI during the early stages of task completion, often before attempting independent thought.

4.2. Correlation Analysis

Table 2

Correlation Matrix among Variables

Variables	1	2	3	4
1. GenAI Use	1			
2. Self-reflection	.42**	1		
3. Learner Autonomy	.36*	.61**	1	

4. Creative Dependency	.48**	-.55**	-.40**	1

Note. *p < .05, **p < .01

The correlation matrix (Table 2) shows significant relationships among the variables. Self-reflection has a strong positive relationship with learner autonomy (r = .61, p < .01), supporting the idea that reflective practices boost independent learning. It is also negatively related to creative dependency (r = −.55, p < .01), indicating that students who reflect more are less likely to rely on GenAI. GenAI use is positively related to both learner autonomy (r = .36, p < .05) and creative dependency (r = .48, p < .01), highlighting its dual effect on learning outcomes.

4.3. Regression Analysis

To further investigate the relationships among variables, multiple regression analyses were performed.

An important observation emerging from the data is that GenAI use was a significant predictor of learner autonomy (β = .32, p < .05), supporting H1. However, its explanatory power was modest (R^2 = .13), indicating that other factors play a more significant role.

Table 3

Regression Analysis: GenAI Use Predicting Learner Autonomy

Variable	B	SE	Beta	t	Sig.
GenAI Use	.28	.11	.32	2.54	.014

R^2 = .13

The results also suggest that self-reflection demonstrated a strong positive effect on learner autonomy (β = .60, p < .001), accounting for a significant portion of the variance (R^2 = .37). This finding is consistent with self-regulated learning theory and supports H2.

Table 4.

Regression Analysis: Self-reflection Predicting Learner Autonomy

Variable	B	SE	Beta	t	Sig.
Self-reflection	.52	.09	.60	5.78	.000

R^2 = .37

To investigate the mediating role of self-reflection (H3), both GenAI use and self-reflection were included in the regression analysis. The results show that self-reflection remained a significant predictor (β = .55, p < .001). At the same time, the effect of GenAI use became non-significant (β = .14, p > .05). This suggests that self-reflection fully mediates the relationship between GenAI use and learner autonomy.

Table 5

Multiple Regression (Mediation Test)

Variable	B	Beta	Sig.
GenAI Use	.12	.14	.210
Self-reflection	.48	.55	.000

R^2 = .41

Finally, regression analysis revealed that self-reflection negatively predicts creative dependency ($\beta = -.52$, $p < .001$), supporting H4. It confirms that reflective learners are less likely to rely on GenAI. Additionally, the findings support H5, indicating that lower levels of self-reflection are associated with increased dependence on AI tools.

Table 6

Regression Analysis: Self-reflection Predicting Creative Dependency

Variable	B	Beta	Sig.
Self-reflection	-.46	-.52	.000

$R^2 = .27$

4.4. Relationship between Reflection and Autonomy

Overall, all proposed hypotheses were supported. The findings confirm that while GenAI use contributes to learning support, its effectiveness depends on students' reflective engagement. Self-reflection emerges as the key factor in enhancing learner autonomy and reducing creative dependency.

4.5. Qualitative findings

The author collected qualitative data from reflective journals and semi-structured interviews, which provided a much clearer sense of how students experienced GenAI. It also helped explain what the quantitative numbers really meant. Three themes emerged from the analysis.

Theme 1: GenAI as a Cognitive Scaffold.

For most of them, GenAI was a support tool. It helped them generate ideas and develop their language, but it did not replace their own thinking. Students often say they used it to get started on a task, consider different viewpoints, or clean up their writing. However, they also pointed out that we have to evaluate and adapt whatever the AI spits out critically. One participant put it like this: "*I usually ask ChatGPT for ideas when I start an essay, but I still need to reorganise and rewrite the content in my own way*" (Student 12). This finding suggests that students generally perceive GenAI as a cognitive scaffold that supports learning while still requiring active intellectual engagement.

Theme 2: Emerging Signs of Creative Dependency

Even students who saw GenAI's benefits had concerns about getting too dependent on it. Their journal entries showed a common worry. Frequent use of AI seemed to chip away at their confidence in generating original ideas on their own. The pattern was clear. They would turn to AI for suggestions before attempting the work themselves. One student was honest about it. "*Sometimes I cannot begin writing until I receive suggestions from AI*" (Student 27). That kind of comment points to something this study calls creative dependency. It describes a gradual shift where AI becomes the primary source of ideas and support. The dependency was probably not absolute for most students. Still, the findings hint at a real risk to independent thinking and to their creative confidence down the line.

Theme 3: Reflection as a Protective Mechanism

The qualitative findings also pointed to how self-reflection helped students manage their use of GenAI. People who took time to think about their interactions with AI seemed more aware of what these tools could and could not do. They tended to review AI suggestions more

carefully, compare them with their own thinking, and decide whether to use them. One student said, "*When I compare my first draft with AI suggestions, I can identify weaknesses in my thinking and improve my work*" (Student 35). So, self-reflection works like a safeguard. It keeps students from relying too heavily on AI by prompting them to reflect on their own thinking. Which means they retain control over their own learning. Moreover, for those like Student 35, that is what makes the difference.

5. Discussion

This study's results mostly back the conceptual framework the author laid out. Self-reflection matters a lot here. It mediates how GenAI use affects learning outcomes. First of all, the data confirmed H1, showing that GenAI use is positively associated with perceived learning support. Students said they often use GenAI to brainstorm ideas, polish their language, and finish assignments. Then, for H2, the results suggest that self-reflection and learner autonomy go hand in hand. Students who reflected more often tended to develop their own ideas and to evaluate the AI's output critically. They also kept their personal voice in their work. More importantly, the study found that the direct contribution of GenAI use became statistically insignificant once self-reflection was introduced into the model. So, this pattern supports H3 in suggesting that reflective engagement may be the mechanism through which learners transform AI-generated assistance into meaningful learning experiences. GenAI shifts from being a tool for efficiency into more of a cognitive partner, one that actually encourages deeper learning. Besides, the results also support H4. There is a clear negative relationship between self-reflection and creative dependence. Students who practised reflection relied less on GenAI and seemed more confident working through tasks independently. Students with low reflection levels, however, often struggled to generate ideas without AI. This action suggests they are gradually depending more on external support than on their own thinking. Finally, the findings support H5: Students who do not reflect much tend to lean more heavily on GenAI for academic tasks. The study calls this creative dependency. Regularly using AI makes independent thinking harder over time. These results confirm the model we proposed. What stands out is how GenAI's impact on learning is not straightforward. It depends heavily on the learner's reflective engagement. Self-reflection is not just a useful habit. It may be what decides whether AI helps a student or makes them more dependent. From a theoretical perspective, the study extends existing models of self-regulated learning by incorporating GenAI as a contextual factor and by emphasising the mediating role of reflection. From a teaching perspective, these findings suggest that educators should go beyond simply integrating AI tools into the classroom and focus on developing students' reflective skills. Without these efforts, widespread use of GenAI could unintentionally reduce learner autonomy and increase dependence. In conclusion, the discussion confirms that the effectiveness of GenAI in English depends on effective thinking, which supports the study's central argument: GenAI supports learning only when it promotes thinking, and reflection is essential for that.

6. Implications

The findings of this study provide significant implications for teaching, curriculum development, and future research in integrating Generative Artificial Intelligence (GenAI) into English Language Studies.

6.1. Pedagogical Implications

The study emphasises the crucial role of self-reflection in improving the effectiveness of GenAI use. Therefore, language teachers should not only promote the use of AI tools but also instruct students in using them thoughtfully. The teachers can achieve this by incorporating

reflective tasks into classroom activities. For example, teachers might ask students to explain how they used GenAI, evaluate its outputs, and compare AI-generated content with their own ideas. Additionally, teachers should assign tasks that require students to demonstrate independent thinking before engaging with GenAI. For instance, students could submit initial drafts or outlines before interacting with AI tools. These practices help prevent passive reliance and encourage deeper cognitive engagement. Furthermore, assessment methods should include reflective components. Instead of solely assessing final products, educators should evaluate the learning process, including students' decision-making and their use of AI. This approach supports the development of learner autonomy and maintains academic integrity.

6.2. Curriculum and Institutional Implications

At the curriculum level, the findings suggest a need to incorporate AI literacy and reflective learning into English language programs. Courses should include explicit instruction on how to critically evaluate AI-generated content, recognise its limitations, and use it responsibly. Institutions should also establish clear guidelines for the proper use of GenAI in academic work. Despite implementing restrictive policies that discourage AI use, universities should adopt a balanced approach that promotes responsible and strategic engagement. Additionally, professional development programs should be offered to help instructors effectively integrate GenAI and reflective practices into their teaching. Teachers may struggle to guide students in meaningful AI use without proper training.

6.3. Theoretical Implications

This study expands the existing literature by broadening the Self-Regulated Learning (SRL) framework to include AI-supported learning environments. The study offers a clearer understanding of how GenAI influences learning outcomes by positioning self-reflection as a mediating variable. Introducing "creative dependency" as a concept enriches the discussion of the unintended effects of AI use in education and offers a useful perspective for future research exploring the balance between technological support and cognitive independence.

6.4. Implications for Future Research

Future research should increase sample sizes and also include participants from diverse institutions and fields to enhance generalizability. Long-term studies are also needed to examine how students' reflective practices and their use of GenAI develop over time. Additionally, developing and validating measurement tools should be focused on assessing reflective AI use across various learning contexts. Experimental studies could also evaluate the effectiveness of specific teaching methods designed to promote self-reflection. In summary, this study highlights that successfully integrating GenAI into education depends on both access to technology and the development of learners' reflective and self-regulatory skills. By fostering these abilities, educators can ensure that GenAI serves as a resource for learning rather than a shortcut that undermines it.

7. Conclusion

This study highlights the importance of self-reflection for effective use of GenAI in English Language Studies. Although GenAI provides significant benefits, its educational value depends on how students interact with it. Besides, self-reflection helps learners remain independent, maintain academic integrity, and deepen their understanding. Therefore, it is essential to incorporate reflective practices in teaching and assessment to maximise GenAI's potential. Future research should involve larger samples and create validated frameworks for measuring reflective AI use across different educational settings.

Reference

Boud, D., Keogh, R., & Walker, D. (1985). *Promoting reflection in learning: A model*. In D. Boud, R. Keogh, & D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 18–40). Routledge.

Chen, X., Zhou, X., & Goh, Y. S. (2026). How Chinese as a foreign language learners use generative AI for oral script-writing: A qualitative perspective on cognitive scaffolding in project-based learning. *Acta Psychologica, 262*, 106071. https://doi.org/10.1016/j.actpsy.2025.106071

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Otto, S., Lavi, R., & Brogaard Bertel, L. (2025). Human-GenAI interaction for active learning in STEM education: State-of-the-art and future directions. *Computers & Education, 239*, 105444. https://doi.org/10.1016/j.compedu.2025.105444

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Rizun, N., Bordean, O. N., Nikiforova, A., Beleiu, I. N., & Revina, A. (2026). Generative AI in higher education: Ethical and behavioural factors influencing students' intentions to use ChatGPT. *Computers and Education Open, 10*, 100336. https://doi.org/10.1016/j.caeo.2026.100336

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Temple Smith.

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, 1261955. https://doi.org/10.3389/fpsyg.2023.1261955

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

TỰ PHẢN TƯ TRONG VIỆC SỬ DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH (GenAI) ĐỂ HỌC TẬP: NHẬN THỨC CỦA SINH VIÊN NGÀNH NGÔN NGỮ ANH TẠI TRƯỜNG ĐẠI HỌC BÌNH DƯƠNG – PHÂN HIỆU CÀ MAU

Liet Thoai Phuong Lan

TÓM TẮT - Trí tuệ nhân tạo tạo sinh (GenAI) đã mang lại những thay đổi căn bản trong giáo dục đại học, đặc biệt là trong lĩnh vực nghiên cứu ngôn ngữ Anh. Tuy nhiên, GenAI chỉ thực sự phát huy hiệu quả tối đa khi đóng vai trò hỗ trợ sinh viên phát triển tư duy độc lập. Do đó, nghiên cứu này tìm hiểu nhận thức và thực hành tự phản tư của sinh viên chuyên ngành tiếng Anh về việc sử dụng GenAI trong quá trình học tập. Thông qua bảng kiểm tự phản tư (self-reflection checklist) có cấu trúc, nghiên cứu khám phá cách sinh viên cân bằng giữa sự hỗ trợ của GenAI và tư duy độc lập trên bốn lĩnh vực chính: giải quyết vấn đề học thuật, biểu đạt sự sáng tạo, giao tiếp và ra quyết định. Nghiên cứu sử dụng phương pháp hỗn hợp (mixed methods), kết hợp bảng hỏi để thu thập dữ liệu định lượng, cùng với nhật ký phản tư và phỏng vấn bán cấu trúc. Đối tượng khảo sát gồm 50 sinh viên chuyên ngành Ngôn ngữ Anh tại Trường Đại học Bình Dương – Phân hiệu Cà Mau. Kết quả sơ bộ cho thấy, dù GenAI được sử dụng rộng rãi để soạn thảo và trau chuốt ý tưởng, việc phát triển tính tự chủ chiến lược là vô cùng cấp thiết nhằm tránh tình trạng rỗng tuếch về sáng tạo và sự lệ thuộc quá mức. Nghiên cứu nhấn mạnh vai trò của tự phản tư như một công cụ sư phạm nhằm thúc đẩy việc tích hợp AI một cách đạo đức và hiệu quả trong đào tạo ngôn ngữ Anh.

Từ khóa: GenAI, tự phản tư, chuyên ngành tiếng Anh, tính tự chủ của người học, liêm chính học thuật

Bionote



MA. Liet, T. P. Lan (Liet Thoai Phuong Lan) earned two master's degrees: an M.A. in English Language from Ho Chi Minh City University of Technology (HUTECH) and an M.A. in Measurement and Evaluation in Education from Vietnam National University, Hanoi. She obtained her bachelor's degree in English from Can Tho University. Since 2000, she has been a teacher at Ca Mau Community College. She currently works as an English language lecturer at Binh Duong University – Ca Mau Campus, Vietnam. With over 25 years of teaching experience in higher education, her academic interests include language teaching methodology, learner autonomy, educational assessment, and the use of emerging technologies, especially Generative AI, in English language learning.

Appendices

STUDENT SELF-REFLECTION ON STRATEGIC GENAI USE

Research Topic: Self-Reflection in Using GenAI for Learning: Perceptions of English-Majored Students at Binh Duong University - Ca Mau Campus

PART 1: GENERAL INFORMATION

Full Name (Optional): _____

Academic Year/Class: _____

Frequency of GenAI use (ChatGPT, Claude, Gemini, etc.):

[ ] Daily [ ] Weekly [ ] Rarely / Never

PART 2: SELF-REFLECTION SURVEY

Please check the boxes that accurately describe your current habits when using GenAI.

Section A: Academic Thinking and Problem-Solving

Strategic Use (Positive Habits):

1. [ ] I attempt problems independently first, then use GenAI to check my approach.
2. [ ] I can explain my work clearly without needing to reference GenAI.
3. [ ] I can complete similar tasks even when GenAI is unavailable.
4. [ ] I immediately turn to GenAI when encountering challenges without trying first.

Section B: Creative Work and Expression

Strategic Use (Positive Habits):

5. [ ] I generate initial ideas independently, then use GenAI to build and explore them.
6. [ ] I maintain a recognisable personal style across all my English assignments.
7. [ ] I feel "creatively empty" without GenAI to start my writing process.
8. [ ] My work has started looking generic or obviously AI-generated.

Section C: Communication and Social Skills

Strategic Use (Positive Habits):

9. [ ] I draft important messages myself, then might use GenAI to refine them.
10. [ ] I can handle unexpected social situations without needing GenAI preparation.
11. [ ] When I need emotional support, I turn to trusted people rather than GenAI.

Section D: Personal Decision-Making

12. [ ] I form my own opinions first, then seek GenAI’s perspective as one input.
13. [ ] My decisions feel authentically mine.
14. [ ] I struggle to make decisions when GenAI is not available.

#61 Exploring EFL Freshmen's Motivation for Choosing a Career in Teaching English: A Case Study at A Public University in Ho Chi Minh City

Son Le Anh, Tai Phan Tan , Thanh Lam Thanh, Tam Huynh Ngoc, Vy Nguyen Pham Tuong, Hieu Duong Thi Le & Ha Nguyen Thi

Abstract

English language teaching is growing rapidly in Vietnam. Therefore, understanding why students choose this career is very important to help universities support and develop future teachers. This study aims to explore the career motivations of freshmen in the English Pedagogy major at a public university in Ho Chi Minh City. It also examines whether gender and previous teaching experience create any differences in their motivations. The study used a quantitative method. Data was collected from 101 first-year students using a 25-item questionnaire based on the FIT-Choice framework. The collected data was then analyzed using SPSS software with descriptive statistics, Pearson correlations, and t-tests. The findings showed that students reported high levels of motivation across all FIT-Choice constructs. Task Demand and Task Return and Socialization Influences received the highest mean scores, while altruistic and intrinsic motivations also emerged as prominent reasons for choosing teaching. Extrinsic motivation, particularly fallback careers, received the lowest ratings. In conclusion, first-year EFL students proactively choose the teaching profession because of their true passion for English and a desire to contribute to society, rather than seeing it as a secondary option. These findings provide helpful evidence for universities to build better strategies to nurture students' passion for teaching.

***Keywords*:** career motivation, FIT-Choice framework, EFL freshmen, pre-service teachers, English Pedagogy

1. Introduction

1.1. Background to the Study

English has become an essential medium for global communication, education, and professional development, leading to an increasing demand for qualified English language teachers in many countries, including Vietnam. As educational systems continue to face challenges related to teacher recruitment and retention, understanding why individuals choose teaching as a career has become an important research concern (Müller et al., 2009; Watt & Richardson, 2007). Career motivation is particularly significant because it influences not only

entry into the profession but also professional commitment, persistence, and long-term career satisfaction (Richardson & Watt, 2006; Watt & Richardson, 2007).

1.2 Statement of the Problem

To investigate motivations for choosing teaching, researchers have widely adopted the Factors Influencing Teaching Choice (FIT-Choice) framework developed by Watt and Richardson (2007). Drawing upon Expectancy-Value Theory (Eccles et al., 1983), the framework explains how individuals' career decisions are shaped by intrinsic interest, social utility values, personal utility values, perceived ability, social influences, and perceptions of the teaching profession. Across diverse educational contexts, studies have consistently found that intrinsic and altruistic motivations are among the strongest reasons for entering the teaching profession, whereas extrinsic considerations such as salary, social status, or fallback

career opportunities tend to be less influential (Asriani et al., 2022; Lee et al., 2019; Pavin Ivanec, 2023; Saito, 2024).

1.3 Research Gap

Within the Vietnamese context, previous research has reported similar patterns, highlighting the importance of personal interest in teaching and the desire to contribute to society (Tran & Cao, 2020). However, recent policy initiatives, particularly Decree No. 116/2020/NĐ-CP, which provides financial support for teacher education students, have introduced additional contextual factors that may shape students' career decisions (Tran et al., 2025). These developments suggest the need for continued investigation of teacher motivation within contemporary Vietnamese educational settings.

1.4 Research Objectives and Research Questions

To fill these gaps, this study aims to explore the career motivations of first-year English pedagogy students at a public university in Ho Chi Minh City. By surveying freshmen without formal teaching experience, this research provides a complete and authentic view of why they chose this major. Specifically, the study seeks to answer the following questions: (1) What are the primary FIT-Choice motivational factors underlying the career choices of first-year English Pedagogy students? and (2) To what extent do gender and prior informal teaching experience differentiate their motivational profiles?

1.5 Significance of the Study

Theoretically, this study extends the application of the FIT-Choice framework to first-year EFL teacher education students, a relatively underexplored population in Vietnam. By focusing on freshmen, the study provides insights into career motivations at an early stage of teacher development before extensive pedagogical training and professional socialization occur. Additionally, the study examines motivational differences across gender and prior teaching experience groups, contributing to a more nuanced understanding of teacher motivation within the Vietnamese context. Practically, the findings offer useful implications for teacher education institutions and policymakers. Understanding the factors that motivate students to pursue English language teaching can support the development of recruitment strategies, orientation programs, and student support initiatives aimed at sustaining motivation and strengthening commitment to the teaching profession.

1.6 Scope and Delimitations

This study focuses on English language teaching first-year students of the Faculty of Foreign Languages at a public university in Ho Chi Minh City, Vietnam. Moreover, data were collected from participants at a single point in time to examine their career motivations. This research intentionally focuses only on first-year English Pedagogy students because they have had limited exposure to pedagogical coursework and teaching practicum, allowing the research to capture their initial career motivations before they are substantially influenced by teacher education. In addition, the study is confined to one public university and adopts a quantitative survey design. Consequently, qualitative perspectives and changes in motivation over time are beyond the scope of the present study.

2. Literature Review

2.1 Definitions of motivation

Motivation is a multifaceted psychological construct that serves as the internal engine for human behavior, initiating and guiding goal-oriented actions through the interplay of

biological, emotional, and cognitive forces (Guay et al., 2010). In its broadest sense, it is defined as the process whereby activity is instigated and sustained (Schunk et al., 2010), varying not only in level but also in orientation, the underlying "why" that moves an individual to act (Deci & Ryan, 2000). In teacher education, motivation is key to attracting and keeping teachers. It decides who joins the profession, how they study during training, and their long-term commitment to teaching (Müller et al., 2009). For first-year students, these early drivers show their "pure" career goals before they face the real challenges of teaching practice (Dewi & Fajri, 2023; Sinclair et al., 2006). Therefore, understanding why students choose this career helps us see how they build their professional identity and stay determined in a university environment.

2.2. Types of motivation

Motivation has been categorized into three main sources: extrinsic, intrinsic and altruistic. Deci & Ryan (2000) demonstrated that motivation includes two kinds: intrinsic and extrinsic. To be more specific, intrinsic motivation refers to doing something because of innate interest or enjoyment, which in the context of teaching may include feelings of enjoyment and satisfaction (Wang & Wang, 2022). On the other hand, extrinsic motivation is related to doing something because of a possible external result (Deci & Ryan, 2000) such as salary (Balyer & Özcan, 2014; Cheung & Yuen, 2016) or the status of the profession (Han & Yin, 2016). Specifically, altruistic motivation refers to the desire to help others and contribute to society (Virtanen et al., 2025). Moreover, Brookhart and Freeman (1992) synthesized research on motivation to conclude that intrinsic, extrinsic and altruistic are the main factors that motivate people to do something, especially choosing a teaching profession, such as making a difference in society and shaping the future of others (Başöz, 2021; Brandmo & Nesje, 2017; Richardson & Watt, 2006). The FIT-Choice framework (Watt & Richardson, 2007) systematically conceptualizes these dimensions, offering a comprehensive model for vocational choice in teaching.

2.3. FIT-Choice Framework

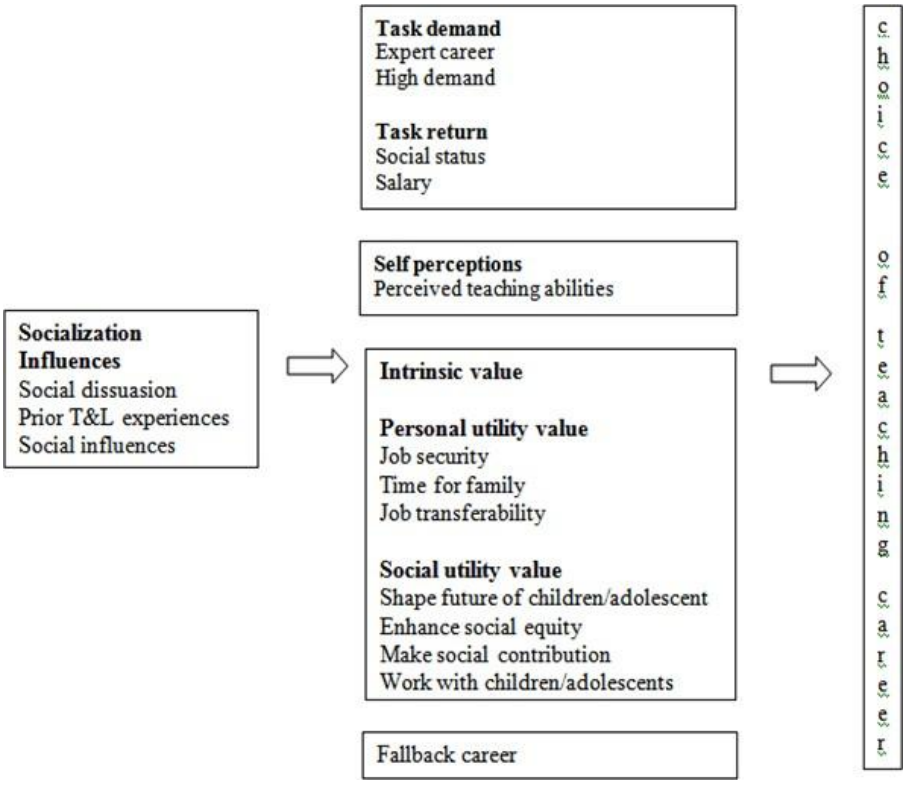
The FIT-Choice model, created by Watt and Richardson (2007), is a popular tool for studying teacher motivation in many countries, including South Korea (Jeong, 2016; Lee et al., 2019), Finland (Goller et al., 2019), Indonesia (Asriani et al., 2022; Sari, 2023), and more recently in the Vietnamese EFL context (Tran & Cao, 2020). Based on the Expectancy-Value Theory, this model suggests that career choices depend on two main things: what a person believes they can do and how much they value the job. Many factors shape this choice, such as past learning experiences and advice from family or friends. A key part of the model is perceived ability, where students judge their own English and teaching skills before joining the profession.

The model also breaks down values into three specific groups. First is intrinsic value, which is the love for the subject or teaching itself. Second is personal utility value, involving job security, time for family and job transferability. Third is social utility value, which is the desire to help children and society. These values are taken into consideration by task perceptions, such as the balance between hard work and rewards like salary or social status.

Finally, the model looks at fallback career factors to see if someone chose teaching just because they failed to get into their favorite major.

Figure 2.3

The Fit-Choice model (Watt & Richardson, 2007, p.176)



Overall, the FIT-Choice framework helps us understand how personal goals, skills, and the desire to help others come together to form a student's professional identity. This provides a strong foundation to explain why they chose English teaching and how committed they will be in the future.

2.4. Previous studies

Global research from Western (Spain, Scotland, Norway, Ireland, Sweden) to Eastern cultures (Indonesia, Korea, Japan), even mixed cultures such as Turkey consistently identifies altruistic and intrinsic factors as the primary motivators for choosing a teaching career. Conversely, extrinsic motivation, particularly fallback careers, salary, and social status, are frequently reported as the lowest-rated motivations (Asriani et al., 2022; Başöz, 2021; Hennessy & Lynch, 2017; Kilinç et al., 2012; Lee et al., 2019; Nesje et al., 2018; Pavin Ivanec, 2023; Saito, 2024; Wang et al., 2022).

In the Vietnamese context, quantitative studies reinforce the dominance of altruistic and intrinsic motivations, often derived from students' prior learning experiences. (Tran & Cao, 2020). However, extrinsic factors have gained prominence, specifically the impact of Decree No. 116/2020/NĐ-CP, which provides significant financial support to teacher candidates (Tran et al., 2025).

2.5. Gap

Although the FIT-Choice framework has been extensively applied in studies of teacher motivation, several gaps remain in the current literature. First, many previous studies have focused on pre-service teachers from mixed year levels or students who have already completed substantial pedagogical coursework and teaching practicum experiences (Tran & Cao, 2020). Such experiences may influence students' professional beliefs and career motivations, making

it difficult to distinguish between initial motivations and motivations that develop during teacher education programs (Sinclair et al., 2006).

Second, relatively limited attention has been given to first-year English Pedagogy students in Vietnam. As freshmen are at the beginning of their academic journey and have had limited exposure to formal teacher training, investigating this population may provide a clearer understanding of the motivations that initially attract students to the teaching profession. Examining these early-stage motivations is important because they may serve as a foundation for later professional identity development and career commitment.

Third, while previous Vietnamese studies have primarily focused on identifying dominant motivational factors, less attention has been paid to potential variations in motivation associated with demographic characteristics such as gender and prior teaching experience. Exploring these differences may contribute to a more nuanced understanding of teacher motivation and provide insights into how different groups of students perceive the teaching profession at the beginning of their university studies.

2.6. Theoretical implications

Using the FIT-Choice framework to study first-year English pedagogy students brings several useful theoretical benefits. First, this study helps expand the FIT-Choice model by testing it on a very specific group: freshmen in a Vietnamese English-teaching context. Because these students have not started real teaching practice yet, looking at their answers helps us understand their original reasons for choosing this job. This adds a new time-based perspective to the current theory, showing what motivates students before they are changed by university training. Second, by looking at how gender and previous experience affect students' choices, the study adds more details to the theory about how personal backgrounds shape a student's values, such as wanting to help society or looking for job security. Finally, this research helps improve the theory on teacher identity. It explains how a student's early belief in their own skills and their desire to help others act as the starting points for their future love and commitment to the teaching career.

3. Methodology

3.1. Research design

This study adopted a quantitative cross-sectional survey design to investigate the career motivations of first-year English pedagogy students at a public university in Ho Chi Minh City, Vietnam. According to Creswell and Creswell (2018), A cross-sectional survey design was considered appropriate because data were collected from participants at a single point in time to examine their career motivations without any intervention or manipulation of variables (Creswell & Creswell, 2018). Furthermore, the research was conducted at the Faculty of Foreign Languages of a public university located in Ho Chi Minh City, Vietnam. With a long history in teacher education, this research location provides an ideal and suitable context to explore what drives freshmen to become English teachers.

3.2. Participants

The participants were carefully selected based on two limiting conditions: they must be first-year students, and they must currently study in the English Pedagogy major. At first, the survey reached 103 students. However, after data cleaning, two invalid responses were removed. Therefore, the final sample consisted of 101 participants. In terms of gender, female students accounted for the majority with 68 participants (67.3%), while male students made up 33 participants (32.7%). Regarding prior teaching experience, 58 students (57.4%) reported

that they had previous teaching experience (such as tutoring), whereas 43 students (42.6%) had no teaching experience.

3.3. Sampling technique

The participants were selected using convenience sampling from the EFL pedagogy classes at a university in Ho Chi Minh. A convenience sampling method was employed because all first-year English Pedagogy classes that were accessible during the data collection period were invited.

3.4. Instruments and Materials

A structured questionnaire was the main tool for collecting data. As noted by Brace (2004), questionnaires are highly effective for gathering a large amount of information about people's attitudes within a specific group. The questionnaire was administered via Google Forms and divided into two main sections:

Section 1: Demographic Information. This section collected basic information, including the students' gender and prior teaching experience.

Section 2: Career Motivations. This section comprised 25 items adapted from the Factors Influencing Teaching Choice (FIT-Choice) framework developed by Watt and Richardson (2007). Minor wording modifications were made to ensure suitability for the Vietnamese higher education context while maintaining the original meanings of the constructs. It measured six main constructs: Intrinsic Motivation (3 items), Altruistic/Social Utility Motivation (4 items), Extrinsic/Personal Utility Motivation (6 items), Perceived Ability (2 items), Task Demand and Task Return (6 items), and Socialization Influences (4 items). All items used a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

Before the main survey, reliability of each construct was subsequently examined using Cronbach's alpha after data collection.

Table 3.4

Reliability statistics for the FIT-Choice constructs

Motivational Constructs	Number of Items	Cronbach's Alpha
Intrinsic Motivation	3	.715
Altruistic/Social Utility Motivation	4	.892
Extrinsic/Personal Utility Motivation	6	.782
Perceived Ability	2	.840
Task Demand and Task Return	6	.850
Socialization Influences	4	.751

3.5. Procedures

Data was collected through four sequential stages and was carried out step by step. First, the questionnaire was adapted from the FIT-Choice framework and modified to fit the Vietnamese context. After that, it was created as an online survey using Google Forms. Before the official collection, a pilot study was conducted in one day with 15 freshmen and 1 lecturer.

The purpose was to check if the questions were clear enough to distribute to the first-year English pedagogy students. Based on feedback from both the lecturer and the pilot participants, the questionnaire was revised and finalized. Next, the official questionnaire was distributed to the first-year English Pedagogy students via Google Forms. Before participating, all volunteers received clear explanations about the study's purpose and provided their consent. This official collection phase lasted for four days. Finally, after the survey had been closed, all responses were screened for completeness and consistency. Incomplete or invalid responses were removed from the dataset. The remaining data were coded and entered into SPSS version 26.0.

3.6. Variables

The study involved two main types of variables. The independent variables were the students' demographic characteristics, specifically their gender (male or female) and prior teaching experience (yes or no). Meanwhile, the dependent variables were the students' career motivations. These motivations were measured through their scores on the 25-item FIT-Choice questionnaire, which covered the six main motivation constructs.

3.7. Data analysis

SPSS software (version 26) was used to analyze the collected data. First, descriptive statistics, including means and standard deviations, were used to summarize the students' motivation levels. Next, a Pearson correlation test was applied to find the relationships between the six motivation factors. Finally, independent samples t-tests were conducted to compare the motivation scores between different groups (male vs. female and experienced vs. inexperienced). These statistical tools provided clear evidence to fully answer the research questions.

4. Results

4.1. Research question 1: What are the primary FIT-Choice motivational factors underlying the career choices of first-year English Pedagogy students?

4.1.1. Overall Picture of Career Motivational Factors

Table 4.1.1

Descriptive statistics of FIT-Choice motivational factors influencing career choice

Motivational Constructs	M	SD
Task Demand and Task Return	4.32	0.61
Socialization Influences	4.14	0.70
Altruistic/Social Utility Motivation	4.10	0.86
Intrinsic Motivation	3.98	0.77
Perceived Ability	3.78	0.85
Extrinsic/Personal Utility Motivation	3.60	0.83

Descriptive statistics indicate that first-year English Pedagogy students reported generally high levels of career motivation across the six FIT-Choice constructs. Among these, Task Demand and Task Return received the highest mean score (M = 4.32, SD = 0.61),

followed by Socialization Influences (M = 4.14, SD = 0.70) and Altruistic/Social Utility Motivation (M = 4.10, SD = 0.86). These findings suggest that participants generally recognized both the characteristics of the teaching profession and the influence of significant others when considering teaching as a career. In contrast, Extrinsic/Personal Utility Motivation received the lowest mean score (M = 3.60, SD = 0.83), although it remained above the midpoint of the scale. This indicates that external or personal utility considerations were less strongly endorsed than other motivational dimensions among the participants. Also, this overall pattern appears inconsistent with the common perception that students choose the teaching profession primarily for extrinsic rewards like the financial support policies from Decree No. 116/2020/ND-CP (Tran et al., 2025) or due to a lack of career orientation. Instead, they enter the field with a clear and realistic awareness of the difficulties and specific characteristics of the teaching profession.

4.1.2. Detailed Analysis of Motivational Constructs

Table 4.1.2

The descriptive statistics of all motivational items based on the FIT-Choice framework

Constructs and Items	M	SD
Task Demand and Task Return		
(E.1) Teaching English requires a high level of specialized knowledge.	4.56	0.73
(E.2) Teaching English requires a high level of pedagogical skills.	4.53	0.74
(E.4) Teaching English is a heavy workload/demanding profession.	4.41	0.80
(E.3) Teaching English requires hard work and high emotional demand.	4.32	0.75
(E.6) English teachers have a good salary/income.	4.21	0.80
(E.5) Teaching is a high-status profession in Vietnam.	3.91	0.98
Socialization Influences		
(F.3) My former teachers have inspired me to be a teacher.	4.28	0.90
(F.4) People around me thought I would make a good teacher.	4.20	0.92
(F.2) My family members influenced my decision to become a teacher.	4.13	0.95
(F.1) My friends influenced my decision to become a teacher.	3.95	0.95
Altruistic/Social Utility Motivation		
(B.2) I want to be a positive influencer on the future of students/pupils.	4.31	0.94
(B.1) I want to contribute to the development of society through teaching.	4.12	0.92
(B.4) Teaching English allows me to benefit the socially disadvantaged.	4.06	0.99
(B.3) I feel pleased because teaching allows me to work with children.	3.92	1.09
Intrinsic Motivation		
(A.1) I love the English language subject and want to work directly with it.	4.24	0.81
(A.2) I believe the teaching profession will bring me long-term enjoyment.	4.03	0.91
(A.3) I have always wanted to become a teacher.	3.68	1.15
Perceived Ability		
(D.1) I have good English language skills for teaching.	3.80	0.91
(D.2) I have good pedagogical skills.	3.76	0.92
Extrinsic & Personal Utility Motivation		
(C.1) Teaching English will be a secure job.	4.12	0.93
(C.6) I think teaching will allow me to work anywhere.	3.93	0.93
(C.3) The teaching profession offers me time for family.	3.80	1.07
(C.2) I do not worry about unemployment when teaching English.	3.65	1.21

Constructs and Items	M	SD
(C.4) I chose this major because of financial support policies (Decree 116).	3.54	1.35
(C.5) I chose this major as a fallback career (secondary option).	2.54	1.56

Table 4.1.2 provides a closer look at the specific survey items, showing some clear trends in why students choose this major. For Task Demand, students showed a very realistic understanding of the job's requirements. They gave very high scores to statements saying that teaching requires a "high level of specialized knowledge" (M = 4.56) and "pedagogical skills" (M = 4.53). These scores are much higher than their expectations for a "good salary" (M = 4.21). These findings indicate that participants generally agreed that teaching requires substantial professional knowledge and pedagogical competence. In addition, among socialization influences, former teachers received the highest mean score among the four socialization items, which is greater than the influence of family (M = 4.13) and friends (M = 3.95) in inspiring their career paths.

Furthermore, their wish to be a "positive influencer" (M = 4.31) and their "love for the English language" (M = 4.24) were much stronger than the thought of "always wanting to become a teacher" (M = 3.68). Even with this strong passion, they rated their English (M = 3.80) and teaching abilities (M = 3.76) at a moderate level. This means they know they need more university training. Finally, although they value job security (M = 4.12), students strongly disagreed with the idea of choosing this major as a "fallback career" (M = 2.54). Because this item has the lowest score, it suggests that their career choices are active and clear, not just a backup plan.

4.1.3. Correlations among Motivational Factors

Table 4.1.3

The correlations between different motivational factors

Constructs	1	2	3	4	5	6
1. Intrinsic Motivation	–	.723**	.494**	.588**	.495**	.541**
2. Altruistic/Social Utility Motivation	.723**	–	.496**	.636**	.495**	.488**
3. Extrinsic/Personal Utility Motivation	.494**	.496**	–	.563**	.420**	.535**
4. Perceived Ability	.588**	.636**	.563**	–	.303**	.417**
5. Task Demand and Task Return	.495**	.495**	.420**	.303**	–	.498**
6. Socialization Influences	.541**	.488**	.535**	.417**	.498**	–

*Note. **p < .01 (2-tailed)*

Pearson correlation analysis revealed that all six motivational constructs within the FIT-Choice model exhibit statistically significant positive correlations with one another (p <.01). This indicates that these motivational factors do not operate in isolation but are intricately intertwined, forming a cohesive psychological ecosystem in the students' career decision-making process.

The most striking finding is the robust positive correlation between Intrinsic Motivation and Altruistic/Social Utility Motivation (r = .723, p < .01). For EFL freshmen, personal passion for the language and the idealistic desire to contribute to society merge into one unified drive. The more they love the English language, the stronger their interest to use it as a tool to help

students and make a positive social impact. Furthermore, Perceived Ability also showed a strong positive relationship with Altruistic Motivation (r = .636) and Intrinsic Motivation (r = .588). This suggests that Students who reported higher perceived ability also tended to report stronger intrinsic and altruistic motivations.

4.2. Research question 2: To what extent do gender and prior informal teaching experience diflerentiate their motivational profiles?

4.2.1. Gender Differences in Career Motivation

Table 4.2.1

Descriptive Statistics of FIT-Choice Motivational Factors by Gender

Motivational Constructs	Male (n = 33)	Female (n = 68)	t	p
Intrinsic Motivation	4.24 (0.71)	3.86 (0.78)	2.403	**.018***
Altruistic/Social Utility Motivation	4.27 (0.85)	4.02 (0.85)	1.343	.182
Extrinsic/Personal Utility Motivation	3.84 (0.89)	3.48 (0.77)	2.104	**.038***
Perceived Ability	4.05 (0.85)	3.65 (0.82)	2.219	**.029***
Task Demand and Task Return	4.45 (0.51)	4.26 (0.65)	1.457	.148
Socialization Influences	4.24 (0.65)	4.09 (0.73)	1.035	.303

*Note. *p < .05*

Independent-samples t-tests were conducted to determine if demographic characteristics, specifically gender, create variations in career motivations. The results indicated statistically significant differences between male and female students in three specific domains.

Male freshmen exhibited significantly higher levels of Intrinsic Motivation (p = .018), Extrinsic/Personal Utility Motivation (p = .038), and Perceived Ability (p = .029) compared to their female counterparts. In a traditionally female-dominated major like English Pedagogy, male students who choose to pursue this path often equip themselves with a stronger intrinsic motivation and higher self-confidence to overcome underlying gender stereotypes. Additionally, driven by traditional societal expectations regarding financial provision, male students also display a more pragmatic concern for extrinsic rewards, such as salary and job security. Nevertheless, both genders share equally high levels of Altruistic Motivation and a realistic understanding of Task Demand (p > .05), proving that the core values of the teaching profession resonate universally across genders.

4.2.2 The Impact of Prior Teaching Experience: Passion and Academic Humility

Table 4.2.2

Descriptive Statistics of FIT-Choice Motivational Factors by Prior Teaching Experience

Motivational Constructs	Male (n = 33)	Female (n = 68)	t	p

Intrinsic Motivation	4.13 (0.69)	3.79 (0.84)	2.202	**.030***
Altruistic/Social Utility Motivation	4.25 (0.75)	3.90 (0.96)	2.056	**.042***
Extrinsic/Personal Utility Motivation	3.66 (0.82)	3.52 (0.84)	.832	.407
Perceived Ability	3.90 (0.72)	3.63 (0.98)	1.588	.115
Task Demand and Task Return	4.45 (0.51)	4.27 (0.63)	.739	.462
Socialization Influences	4.14 (0.66)	4.13 (0.76)	.060	.952

Note. $*p < .05$

Independent samples t-tests indicated significant differences based on prior teaching experience. Freshmen with prior experience demonstrated significantly higher Intrinsic Motivation ($p = .030$) and Altruistic Motivation ($p = .042$) than those without such experience. Initial interactions with learners successfully ignited their love for teaching and magnified their desire to become inspiring educators.

Most importantly, this practical experience did not foster complacency or an illusion of mastery. There was no significant difference in Perceived Ability ($p = .115$) between the experienced and inexperienced groups. This highlights a remarkable sense of academic humility among the freshmen. Despite having prior tutoring experience, they accurately perceive the vast gap between informal tutoring and professional classroom teaching. Consequently, they maintain a modest attitude, recognizing the necessity for rigorous, formal pedagogical training at the university level.

5. Discussion

5.1. Purpose

This study explored the career motivations of first-year English Pedagogy students at a public university in Ho Chi Minh City, while examining the influence of gender and prior teaching experience on these drives.

5.2. Findings

Although Task Demand and Task Return and Socialization Influences obtained the highest overall scores, the item-level analysis indicates that students were primarily attracted to teaching because of altruistic and intrinsic values rather than external rewards. This aligns with Tran and Cao (2020), emphasizing the role of passion and social contribution. It also indicates that modern EFL students actively and proudly choose this path. Interestingly, the results disagree with Tran et al. (2025), as Decree No.116 did not strongly influence their career choice, highlighting the students' genuine devotion to the profession. Demographically, male students showed significantly higher Intrinsic and Extrinsic motivations, as well as Perceived Ability. Moreover, freshmen with prior teaching experience had stronger passion, but their Perceived Ability stayed at a realistic, moderate level compared to inexperienced students.

5.3. Explanation for findings

These outcomes may be associated with the deep-seated cultural respect for the teaching profession. The high levels of Altruistic and Intrinsic motivation, together with participants' awareness of task demands, may indicate that students have some understanding of the profession's requirements rather than viewing it solely through an idealized lens. This pattern may suggest a more nuanced understanding of the profession, although further research would be needed to explore how such perceptions develop over time. Regarding gender, higher

motivation among males likely stems from the need to overcome stereotypes in a female-dominated major (Tran & Cao, 2020; Tran et al., 2025). Their extrinsic drive may be associated with societal expectations for financial stability. Finally, the impact of prior experience illustrates an increased awareness of professional knowledge; while tutoring fuels passion, students recognize the necessity of formal university training to bridge the gap between informal instruction and professional teaching.

5.4. Limitations

Despite providing valuable insights, this study has a few limitations. First, the self-reported data with small sample size (N = 101) from a single public university in Ho Chi Minh City means the findings should be interpreted with caution and may not be generalizable to all EFL teacher education students in Vietnam. Second, using a self-reported questionnaire can cause social desirability bias, as participants might unconsciously rate altruistic items higher to fit the noble image of a teacher. Lastly, the cross-sectional quantitative design only captures a snapshot of their current motivations and lacks the deep qualitative data needed to fully explain the psychological processes behind these choices.

5.5. Implications

Universities should introduce early practical opportunities, such as tutoring clubs, to sustain teaching passion. Faculties must support male students through mentorship programs and career talks featuring successful male educators. Future research should utilize qualitative or longitudinal methods to track how motivations evolve throughout the four-year curriculum.

Theoretically, the findings may contribute to the growing body of research applying the FIT-Choice framework in EFL teacher education contexts. The results also point to the potential influence of cultural and gender-related factors on career choice motivations, which could be explored further to enhance current theoretical explanations of why students choose teaching as a career.

5.6. Recommendations

Institutions should prioritize early-exposure initiatives to maintain freshmen's enthusiasm. Recruitment campaigns should emphasize social impact and personal fulfillment over financial incentives. Additionally, establishing mentorship for male students is crucial for gender inclusivity. Finally, freshmen should be encouraged to engage in informal teaching early to build a grounded professional identity and ensure long-term commitment.

6. Conclusion

This study investigated the initial career motivations of first-year English Pedagogy students at a public university in Ho Chi Minh City using the FIT-Choice framework, alongside the impacts of gender and prior teaching experience.

Findings reveal that freshmen are primarily driven by Altruistic and Intrinsic motivations, decisively rejecting teaching as a fallback career. Notably, male students exhibited higher motivation, potentially to overcome gender stereotypes, while those with prior tutoring experience demonstrated academic humility by recognizing the gap between informal tutoring and professional teaching.

However, generalizability is limited by a small sample size and potential social desirability bias. Future research should employ qualitative and longitudinal methods to track how these motivations evolve throughout the students' four-year academic curriculum and formal teaching practices.

References

Asriani, I. E., Apriliaswati, R., & Riyanti, D. (2022). Motivational Factors Influencing Pre-Service EFL Teachers to Choose Teaching as A Future Career. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, *9*, 38. https://doi.org/10.33394/jo-elt.v9i1.5132

Balyer, A., & Özcan, K. (2014). Choosing teaching profession as a career: Students' reasons. *International Education Studies*, *7*, 104–115. https://doi.org/10.5539/ies.v7n5p104

Başöz, T. (2021). "I've Decided to Become an English Teacher": Pre-Service EFL Teachers' Reasons for Choosing Teaching as a Career. *Kastamonu Eğitim Dergisi*, *29*, 813–821. https://doi.org/10.24106/kefdergi.860727

Brace, I. (2004). *Questionnaire design: How to plan, structure and write survey material for effective market research.* www.kogan-page.co.uk

Brandmo, C., & Nesje, K. (2017). Factors motivating students to become secondary school teachers evidence from Norway. In *Global Perspectives on Teacher Motivation* (pp. 95–125). Cambridge University Press. https://doi.org/10.1017/9781316225202.004

Cheung, A. C. K., & Yuen, T. W. W. (2015). Examining the motives and the future career intentions of mainland Chinese pre-service teachers in Hong Kong. *Higher Education*, *71*, 209–229. https://doi.org/10.1007/s10734-015-9897-3

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Dewi, S. S., & Fajri, M. S. Al. (2023). Investigating Pre-Service EFL Teachers' Initial Motivations and Identity Formation: Evidence From an Indonesian University. *SAGE Open*, *13*. https://doi.org/10.1177/21582440231209123

Goller, M., Ursin, J., Vähäsantanen, K., Festner, D., & Harteis, C. (2019). Finnish and German student teachers' motivations for choosing teaching as a career. The first application of the FIT-Choice scale in Finland. *Teaching and Teacher Education*, *85*, 235–248. https://doi.org/10.1016/j.tate.2019.06.023

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H., Larose, S., & Boivin, M. (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *British Journal of Educational Psychology*, *80*, 711–735. https://doi.org/10.1348/000709910X499084

Han, J., & Yin, H. (2016). Teacher motivation: Definition, research development and implications for teachers. *Cogent Education* (Vol. 3). Taylor and Francis Ltd. https://doi.org/10.1080/2331186X.2016.1217819

Hennessy, J., & Lynch, R. (2017). "I chose to become a teacher because". Exploring the factors influencing teaching choice amongst pre-service teachers in Ireland. *Asia-Pacific Journal of Teacher Education*, *45*, 106–125. https://doi.org/10.1080/1359866X.2016.1183188

Eccles, J. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motives: Psychological and sociological approaches (pp. 75-146). San Francisco, CA: Free man.

Jeong, K. O. (2016). A Study on Korean EFL Pre-Service Teachers' Motivations in Choosing Teaching as a Career. *2016 International Conference on Platform Technology and Service, PlatCon 2016 - Proceedings*. https://doi.org/10.1109/PlatCon.2016.7456779

Kilinç, A., Watt, H. M. G., & Richardson, P. W. (2012). Factors Influencing Teaching Choice in Turkey. *Asia-Pacific Journal of Teacher Education*, *40*, 199–226. https://doi.org/10.1080/1359866X.2012.700048

Lee, J. A., Kang, M. O., & Park, B. J. (2019). Factors influencing choosing teaching as a career: South Korean preservice teachers. *Asia Pacific Education Review*, *20*, 467–488. https://doi.org/10.1007/s12564-019-09579-z

Müller, K., Alliata, R., & Benninghoff, F. (2009). Attracting and Retaining Teachers: A Question of Motivation. *Educational Management Administration and Leadership*, *37*, 574–599. https://doi.org/10.1177/1741143209339651

Nesje, K., Brandmo, C., & Berger, J. L. (2018). Motivation to Become a Teacher: a Norwegian Validation of the Factors Influencing Teaching Choice Scale. *Scandinavian Journal of Educational Research*, *62*, 813–831. https://doi.org/10.1080/00313831.2017.1306804

Pavin Ivanec, T. (2023). Motivation for the choice of a teaching career: comparison of different types of prospective teachers in Croatia. *Journal of Education for Teaching*, *49*, 355–369. https://doi.org/10.1080/02607476.2022.2113736

Richardson, P. W., & Watt, H. M. G. (2006). Who Chooses Teaching and Why? Profiling Characteristics and Motivations Across Three Australian Universities. *Asia-Pacific Journal of Teacher Education*, *34*, 27–56. https://doi.org/10.1080/13598660500480290

Saito, A. (2024). "I chose to receive teacher training because …" Motivations for teaching and career aspirations among teacher education students in Japan. *International Journal of Educational Research Open*, *6*. https://doi.org/10.1016/j.ijedro.2024.100330

Sari, M. A. D. (2023). *Motivational Factors of Indonesian Pre-service Teachers Influencing Teaching As A Career Choice: A Survey Study*. Https://Dspace.Uii.Ac.Id/Handle/123456789/57657;Jsessionid=A2D78621898 7F1B67714938AEC45B8D0.

Schunk, D. H., Pintrich, P. R. & Meece, J. L. (2010). Motivation in Education: Theory, Research, and Applications (3th ed.).

Sinadia, A. R., Kondo, P., & Sunra, L. (2024). Freshmen's Motivations and Career Objectives for Taking the Major of English Education. *Journal of English Language and Education*, *9*, 170–179. https://doi.org/10.31004/jele.v9i6.604

Sinclair, C., Dowson, M., & McInerney, D. M. (2006). Motivations to teach: Psychometric perspectives across the first semester of teacher education. *Teachers College Record* (Vol. 108, pp. 1132–1154). Teachers College, Columbia University. https://doi.org/10.1111/j.1467-9620.2006.00688.x

Tran, T. H. D., & Cao, T. H. (2020). Inside an English Language Teacher Education Program in Vietnam: Students' Motivations for Teaching and Their Intentions to Teach. *VNU Journal of Foreign Studies*, *36*. https://doi.org/10.25073/2525-2445/vnufs.4555

Tran, T. K. A., Phan, K. T., & Tran, T. T. Q. (2025). The Impact of Decree No. 116/2020/NĐ-CP on First-Year Students' Decisions to Choose Teaching Majors at Dalat University. *Vietnam Journal of Educational Sciences*, *21*(S3). https://doi.org/10.15625/2615-8957/12520316

Virtanen, P., Watt, H. M., Richardson, P. W., Linnanmäki, N., Mankki, V., Metsäpelto, R. L., & Wallinheimo, K. (2025). Finnish teacher students' career choice motivations: A mixed methods study. *Teaching and Teacher Education*, *163*, 105068. https://doi.org/10.1016/j.tate.2025.105068

Wang, W., & Wang, Z. (2022). Why choose a career in teaching? Exploring motivational factors that influence the decision to teach. *British Journal of Guidance and Counselling*, *51*, 46–57. https://doi.org/10.1080/03069885.2022.2040003

Wang, W., Wang, Z., & Lin, W. (2022). 'I really like teaching, but…' understanding pre-service teachers' motivations and perceptions about teaching as a career choice. *Journal of Education for Teaching*, *49*, 194–206. https://doi.org/10.1080/02607476.2022.2061337

Watt, H. M. G., & Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: Development and validation of the FIT-choice scale. *Journal of Experimental Education*, *75*, 167–202. https://doi.org/10.3200/JEXE.75.3.167-202

#63 Vietnamese Students' Readiness for English as a Second Language: The Role of AI-Supported Learning Environments

Vu Thanh Loan[1]

Abstract

With the trend of English as a second language (ESL) as a major priority in Vietnam, students in universities are expected to use English in addition to the classroom more frequently than before. However, in a country that encourages widespread adoption of digital learning tools, this phenomenon may remain a hurdle as they tend to have difficulty in real-life ESL applications, leading to a significant discrepancy between what they learn and actual language use. Most previous studies have primarily focused on learners' motivation, perceptions, and attitudes toward AI-assisted language learning, while relatively little attention has been paid to the mechanisms through which AI-supported learning environments influence students' readiness to function in ESL-oriented contexts. The purpose of this study was to examine Vietnamese university students' ESL readiness and explore how AI-supported learning environments contribute to the development of such readiness through learner engagement and self-directed learning. A mixed-methods design was employed, incorporating survey data available from 70 university students as well as semi-structured interviews with 10 participants. Student engagement, self-directed learning, and use of English beyond the classroom were examined in this study. Quantitative data was analyzed using descriptive and inferential statistics and qualitative data were analysed thematically. A report on this survey found that this AI-driven learning experience increased student engagement and increased real-life English use, but the communicative competencies of the same students varied. This indicates that once implemented into pedagogy, AI can promote ESL readiness. This research needs context-sensitive utilization of AI in Vietnamese higher education.

***Keywords*:** AI-supported learning; ESL readiness; student engagement; real-world language use; Vietnamese higher education.

1. Introduction

Over the past few years, the country has been transitioning from learning English as a Foreign Language (EFL) to learning English as a Second Language (ESL), in which learners are learning not only linguistic concepts, but also how to speak English for the purposes of communicating in academic, professional, and social settings. This shift puts ever more pressures on learners' communicative effectiveness and their ability to interact with English extracurricularly. On the other hand, in many cases, Vietnamese university students still face challenges to use knowledge learned in teaching and writing classroom applications, indicating a disconnect between a language learning context and the real-life application of that language. However, with the exponential emergence of Artificial Intelligence (AI), the prospects of language learning have been opened up. For instance, chatbots, automated feedback systems, and adaptive learning platforms that utilize AI support have been demonstrated to foster learner engagement, autonomy, and self-regulated learning behaviors (Pan et al., 2024; Yang, n.d.). Studies show that AI can assist with language skill learning, as well as enhance learners' motivation and satisfaction (Shah et al., 2026; Thanh, 2025). In the Vietnamese case, learners show favourable attitudes toward AI-integrated language learning, especially related to enhanced speaking skills and discipline-specific English use (Pham et al., 2024; Quang &

[1] *Vu Thanh Loan, English Department, FPT University, Hanoi, Vietnam; Email:loanvt4@fe.edu.vn*

Minh, 2025). However, despite these encouraging findings, existing studies have predominantly examined learners' perceptions, technology acceptance, and discrete language-skill development, with limited attention to how AI-supported learning environments contribute to students' readiness to engage in authentic English use beyond classroom settings. In the present study, ESL readiness is conceptualized as a multidimensional construct encompassing learners' communicative confidence, self-directed learning capacity, willingness to communicate, and ability to use English meaningfully in authentic contexts. Additionally, the relationship of AI use, learner engagement and communicative competence has not been well investigated (Vistari & Yuliasri, 2026). To fill this gap, the current study (1) seeks to know the preparedness level of Vietnamese college students for L2 learning, (2) the influence of AI supported learning environment on English acquisition readiness, and (3) how use of AI relates to student engagement and language use at the level of the world language. The study aims to respond to the below research questions:

(1) How prepared does Vietnamese university students perceive themselves to be for English as a second language?

(2) How are AI learning environments contributing to the students' engagement and ESL readiness?

(3) What is the link between AI use and students' use of English outside the classroom? A further contribution of the study is finding out about how AI can assist the move from English as Foreign Language (EFL) toward English as Second Language (ESL) as language teaching in the Vietnamese higher education context.

2. Literature Review

2.1. Language Learning with the Assistance of AI in EFL/ESL Contexts

Artificial Intelligence (AI) was integrated into language education in the early 2020s, it altered the interaction between learners, content, teachers, and learning environment. AI-driven devices such as chatbots, automated writing appraisals, and adaptive learning systems offer immediate feedback, personalized learning trajectories, and more practice opportunities. These are the features that have been found to stimulate students' attention and to increase their involvement in language learning tasks (Pan et al., 2024; Wang et al., 2022). Indeed, AI tools can also simulate situations of authentic communication making language use practice in low-risk settings (which is particularly helpful in contexts where exposure to real language use is restricted). For example, it has also been shown through empirical studies that in the case of AI, it may be used to support some specific language learning activities. For instance, AI-powered applications have been found to enhance speaking skills by providing real-time feedback and interactive conversation practice (Giang et al., 2025). Similarly, AI-supported teaching can enrich reading understanding and self-controlled learning by scaffolding learners' participation in structured as well as adaptive tasks (Pan et al., 2024). These results imply that AI is not an additive but a transformative influence on language learning experiences. Engage learners, give autonomy, and integrate AI. An increasing amount of studies explore the potential of AI in promoting learner engagement and autonomy. Self-directed learning AI enabled environments sometimes promote self-directed learning, since students gain the ability to shape the speed and content as well as the mode of their activities. Such a pattern is consistent with research that is associated with the development of self-control, autonomy & independent learners via AI (Yang, n.d.). In addition, learners experience AI tools to be available to them and to provide support, which can motivate and enhance their motivation and willingness to perform language learning tasks (Alkandari, 2025). But engagement with AI-accommodated

contexts is not ubiquitous for learners. Although there exists an active use of such tools by some students to expand their vocabulary outside the classroom, there is also passive or limited use among the students. Based on a study conducted by Vistari and Yuliasri (2026) suggests that learners' attitudes mediate the association between AI use and communicative confidence, suggesting that positive attitudes may not automatically lead to language development. The findings underscore the need for examining not only whether students ultilize AI, but how students use it, in relation to language learning, and language outcomes.

2.2. AI in the Vietnamese Context

In recent years, there has been an increasing prevalence of AI-based language learning in Vietnam. The vast majority of university students accept the use of AI-based tools to enhance the usage of both of these and the perception of the positive effects on English as an EFL skill and learning efficiency is evidenced by previous research (Pham et al., 2024). It has been found that AI applications enhance learner satisfaction and perceived teaching usefulness in language learning environments (Thanh, 2025). Additionally, students had mentioned using AI in support of ESP, which implied that use of AI can assist language learning in a more focused and context-based and targeted way, thus it may assist ESL (Quang & Minh, 2025). However, such positive developments notwithstanding, issues persist in the successful application of AI applications to educational contexts. Problems like over-reliance and limitations in critical usage of, as well as access to, the AI, too much trust by students as well as a lack of access to the technology may hinder its effects. Additionally, the majority of other research in the Vietnamese context prior to that period has concentrated solely on students' perception and intention for learning in a Vietnamese context and less on learning behaviour or learning as an observable outcome.

2.3. ESL Readiness and Research Gap

The idea of ESL readiness lies not only in learners' attitudes/ motivation, but also in how easily they can use English in the real world. This entails a blending of communicative skills, learner control, and ability to interact with English in a range of contexts. Although AI has been proven to support multiple facets of language learning, their role in advancing ESL readiness is less researched, especially in EFL-to-ESL transitioning environments like Vietnam. Moreover, prior work has not explored interaction of the use of AI, student engagement, and real-world language use as interrelated determinants of ESL readiness. This gap underscores the importance of a more rich study that takes into account the quantitative and qualitative aspects of students' experiences within the learning process. Beyond examining ESL readiness as an outcome variable, this study proposes that learner engagement serves as a critical explanatory mechanism through which AI-supported learning environments influence readiness for authentic language use. While previous studies have largely investigated direct relationships between AI adoption and language learning outcomes, the present study extends the literature by positioning engagement as a mediating construct linking technology use with communicative readiness. This perspective contributes to a more process-oriented understanding of AI-assisted language learning in EFL-to-ESL transition contexts.

2.4. Theoretical Framework

The integrated theoretical frameworks that support this study are Self-Determination Theory (SDT), Community of Inquiry (CoI) framework, and ESL readiness. In tandem, these viewpoints offer a basis for interpreting how AI-assisted learning environments are likely to play a role in enhancing students' readiness for English as a second language.

First, Self-Determination Theory (Deci & Ryan, 2000) indicates that learners are motivated and more engaged when their needs for autonomy, competence, and relatedness are satisfied. The study will investigate AI-facilitated language learning tools that can promote learner autonomy or provide on-demand feedback and support whenever students prefer, to ensure they are able to study at their own pace. Such features can enhance the feeling of competence and motivation of the learners toward language learning, which can drive more language learning.

Second, in the Community of Inquiry framework (Garrison et al., 2000), meaningful learning happens through the interplay between cognitive presence, social presence, and teaching presence. In AI-enabled learning environments, AI tools can support a cognitive presence by giving explanations, feedback, and chances to practice. Moreover, teachers are critical in shaping learning environments, which may encourage students to move into meaningful language use. On this view, AI acts as a facilitator, facilitating engagement instead of replacing human interaction.

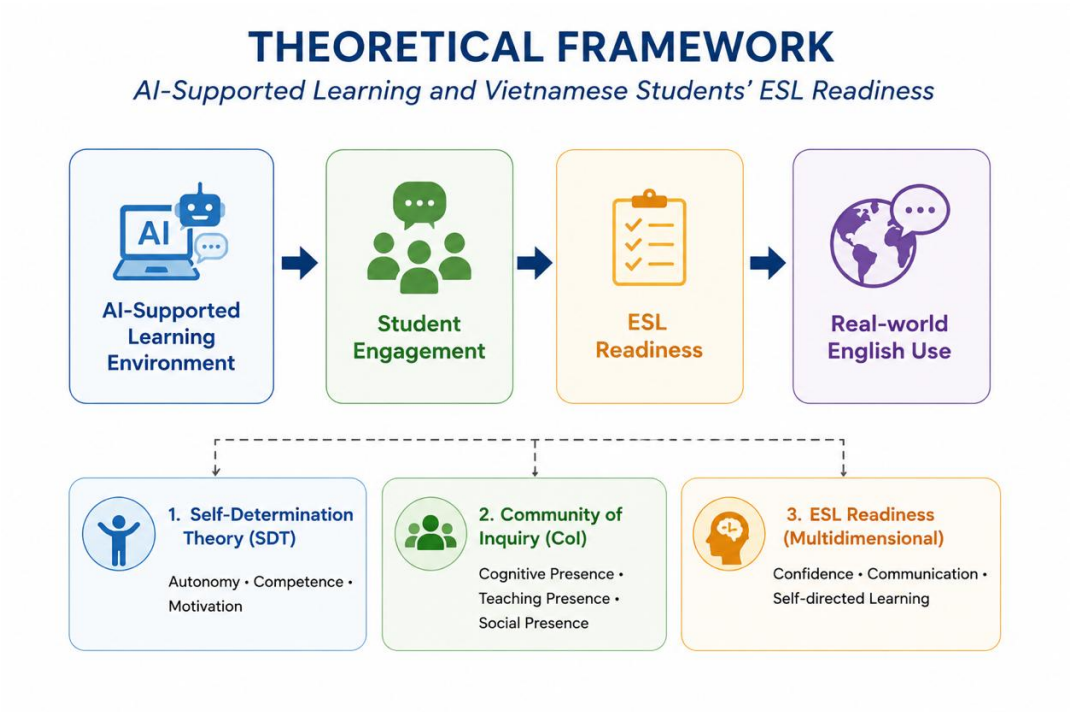
Third, ESL readiness is not limited to language proficiency, but consists of learners' confidence, willingness to communicate, self-directedness to achieve their goals or use English in meaningful contexts. Readiness is accordingly a multidimensional construct indicating linguistic and behavioral readiness for English in the real world. Based on these theoretical constructs, the study argues that AI-enabled learning ecosystems will boost student engagement by promoting autonomy, capability, and cognitive input. That further increases engagement in turn contributes to improved ESL readiness and more frequent, meaningful use of English beyond the classroom. Unlike previous studies that primarily examined direct effects of AI use on language-learning outcomes, the present framework conceptualizes student engagement as a mediating mechanism that explains how AI-supported learning environments may facilitate ESL readiness and authentic language use. By emphasizing the process through which AI influences learning behaviours, the framework offers a more nuanced explanation of AI-assisted language learning outcomes. Therefore, this research proposes the relationship between the main variables as follows:

AI-Supported Learning Environment → Student Engagement → ESL Readiness → Real-world English Use

This framework guided the structure of our research questions, and both the quantitative and qualitative findings were interpreted accordingly.

Figure 1

Theoretical Framework



3. Methodology

3.1. Research Design

This study employs a mixed-methods design in order to investigate Vietnamese university students' English as a second language (ESL) readiness and the extent to which learning environments with AI support impact this. The combination of quantitative and qualitative methodologies allowed the study to identify generalizable trends as well as nuances of students' experiences. In particular, the quantitative one investigated relationships among relevant variables, while the qualitative one offered insights into how and why learners are using AI in language acquisition.

The study included 70 undergraduate students enrolled in English language courses at a private university in Vietnam. Participants took advantage of convenience sampling owing to accessibility and feasibility reasons. This was a very wide variety of students, with varying degrees of proficiency in language and learning style. As for the qualitative research, 10 participants were purposively sampled for the qualitative phase on the basis of their reported prevalence of AI usage (high users, medium users, and low users). This sampling approach facilitated a more balanced examination of different learner experiences and attitudes towards AI-driven language acquisition.

3.2. Instruments

Questionnaire

A standardized questionnaire was developed from the previous available literature on AI-supported language learning and ESL readiness. There were four main sections to the instrument:

Demographic Information (e.g., gender, major, years of learning English)

AI Usage (the frequency and purposes of using AI tools, such as chatbots, writing assistants, and translation tools)

Learner Engagement and Self-directed Learning (behavioral, emotional, and cognitive engagement)
ESL Readiness (confidence in communication, willingness to use English, and frequency of English use beyond the classroom)

Items were assessed using a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). Two language education specialists reviewed content validity for the questionnaire.

Reliability Analysis

Cronbach's alpha coefficients were calculated for each construct to evaluate internal consistency. Results suggested acceptable to high reliability:

Table 1

Reliability of Scales (Cronbach's Alpha)

Construct	Number of Items	Cronbach's Alpha
AI Usage	5	0.81
Engagement	6	0.86
Self-directed Learning	5	0.83
ESL Readiness	6	0.88

All values of the measurement scales were above the minimum recommended value of 0.70, indicating good internal consistency and good reliability.

Interviews

Semi-structured interviews were carried out to probe in more depth into students' perceptions and experiences. A questionnaire that included the following questions was drawn up for the examination:

- How students can use AI tools in learning English?
- Perceptions of benefits and challenges to AI-driven learning?
- Influence of AI on engagement and confidence?
- Real-life experience of using English?

Data Collection Procedures

The data was collected during a 4-week period. The questionnaire was distributed via Google Forms and was voluntary. The participants were informed of the purpose of the study; the process of data collection was preceded by confidentiality and anonymity. Ten participants were invited for follow-up interviews following the survey.

Data Collection and Analysis

Analysis of all collected and coded data was carried out independently of prior research. The interviews (e.g., through Zoom, Google Meet) lasted approximately 20–30 minutes and were audio-recorded with consent of participants. The interviews were subsequently transcribed for analysis.

Quantitative Analysis

Quantitative data were analyzed with statistical software. Descriptive statistics (means and standard deviations) were used to summarize students' responses. By doing correlation

tests using Pearson correlation, inferential analysis was conducted to investigate relationships between AI use, engagement, and ESL readiness.

Qualitative Research Analysis

Qualitative analysis was conducted. Thematic analysis method was used to analyze the qualitative data. The analysis was performed using a systematic way:

(1) Familiarization with the data.

(2) Initial coding.

(3) Identification of themes and sub-themes.

(4) Theme refinement and interpretation. To lend credibility across themes, recurring patterns were compared and representative descriptions were used to reinforce each theme's credibility.

Ethical Considerations

The study was voluntary, and informed consent was given to all participants. The data were anonymized for confidentiality, and the right of the participants to withdraw at any time during the study was informed to them as well.

4. Results

4.1. Quantitative Findings

4.1.1. Descriptive Statistics

Table 2

Descriptive Statistics of Key Variables (N = 70)

Variable	Mean	SD	Level Interpretation
AI Usage Frequency	3.82	0.71	Moderately high
Student Engagement	3.76	0.65	Moderately high
Self-directed Learning	3.69	0.68	Moderate
ESL Readiness	3.58	0.72	Moderate
Real-world English Use	3.41	0.74	Moderately low

Descriptive statistics showed a stark difference in technology engagement vs actual language readiness. AI usage (M = 3.82) and engagement (M = 3.76) score relatively high, but ESL readiness (M = 3.58) and real-world English use (M = 3.41) lag behind. It indicates that regular use of AI tools does not imply proficiency in a foreign language. The gap of approximately 0.4 points between AI usage and real-world use points to a disconnect between digital practice and real-world application. And again, this is a classic example of an issue in EFL contexts that is transitioning to ESL, where one is exposed to the input language but without sufficient practice or confidence to use the language in situations outside of controlled situations. Also, the moderate level of self-directed learning (M = 3.69) indicates that AI tools offer independence while the students might not take the full advantage of these tools for further study. AIs may also serve task completion instead of the skill enhancement purpose which would reduce its impact at ESL readiness.

4.1.2. Correlation Analysis

Table 2

Correlations Between Key Variables

Variables	1	2	3	4
1. AI Usage	—			
2. Engagement	.52**	—		
3. ESL Readiness	.47**	.61**	—	
4. Real-world English Use	.39**	.55**	.63**	—

Note: p < .01

Correlation analysis sheds light on the relationships among variables. AI usage has a moderate correlation with engagement (r = .52) and ESL readiness (r = .47), suggesting AI plays a positive but not strong enough role to affect readiness in a direct manner. Interestingly, engagement has the most significant relation with ESL readiness (r = .61) and use of English in the real world (r = .55). As such, engagement seems to be an important mediating variable, connecting AI usage with language outcomes. In short, AI is effective only when it provokes learners to engage with its meaning. ESL readiness correlates with real-world English use most strongly (r = .63), indicating that readiness is related to real-world use of the language. However, the weaker correlation between AI usage and real-world use (r = .39) points to the fact that AI alone cannot foster a real use of language. Together, these results imply a three-layer relationship:

- AI leads to better engagement.
- Engagement enhances readiness.
- Readiness leads to real-world use.

The complex nature of this relationship supports the assertion that pedagogical mediation is vital to unlock AI's potential in language learning.

4.2. Qualitative Findings

Thematic analysis yielded three major themes with clearly defined sub-themes:

Table 3

Thematic Analysis of Students' Experiences with AI-Supported Learning

Theme	Sub-themes	Description	Representative Insights
1. AI as a Cognitive Support Tool	Instant feedback	AI provides quick corrections and explanations	Students rely on AI for immediate answers
	Vocabulary expansion	Exposure to new words and expressions	AI helps generate synonyms and ideas
	Task assistance	Support for writing and assignments	AI is used to complete tasks efficiently
2. AI-Driven Engagement and Confidence	Increased interest	Learning becomes more interactive and enjoyable	AI makes lessons less monotonous
	Reduced anxiety	Safe environment for practice	Students feel less afraid of making mistakes
	Perceived autonomy	Greater control over learning pace	Learners can study independently

3. Limited Transfer to Authentic Use	Passive dependence	Over-reliance on AI-generated responses	Students copy rather than produce language
	Lack of real interaction	Limited exposure to human communication	AI cannot fully simulate real conversations
	Low communicative confidence	Hesitation in real-life situations	Students feel unprepared outside class

Qualitative findings show that students mostly used AI as a cognitive support factor for making language learning more efficient. They frequently said they worked with such AI to proofread vocabulary, hone grammar and brainstorm ideas for assignments. This would lend credence to the idea that AI is a reinforcement of language input and task achievement, but more thorough analysis reveals a crucial trade-off in efficiency versus deep learning. Many students reported reliance on AI-generated responses without fully applying the language, suggesting a surface-level mode of engagement. This sheds light on the reason why frequently using AI did not always correlate with an increase in ESL readiness, that is, because less cognitive effort reduces the amount of time that learners are needed to sustain use and learn languages independently in the long haul. Besides cognitive assistance, AI was also used as a significant influence in forming the students' affective engagement. The learners felt more motivated and less anxious when responding to AI, largely as a result of a lack of judgment and social influence. The low-risk environment promotes frequent use of English and the feeling of working independently within the framework of learning. But this kind of participation seemed more affective than cognitive. Although they felt comfortable using English with AI, students were not routinely encouraged to think and act critically or make decisions on their own. Consequently the interaction generated by AI remained context-bound and did not fully equip learners with what was required for authentic communication. A less-obvious finding of it from the data is the scant link between AI-assisted learning and genuinely using language. Students noted a reliance on AI for responding, and also less confidence when they were expected to speak on their own. This indicates the emergence of what can be termed "assisted competence" where learners are effective on devices but lose in an unmediated environment. In AI-mediated contexts, the study lacks authentic interaction, negotiation of meaning, and real-time communicative competence, which is critical to ESL readiness. Collectively, the qualitative studies show that while AI does help to increase accessibility, motivation and learning efficiency, this does not lead to proper language use. How effectively AI serves to contribute to this shift from EFL to ESL depends upon the extent to which it integrates in pedagogies focused on active production, critical involvement as well as authentic communication with the learner.

5. Discussion

This research sought to reveal some insights about the readiness of Vietnamese university students towards English as a second language (ESL) and how AI-assisted learning experiences may contribute to this readiness. Findings contribute to provide considerable knowledge on the impact of AI on language learning experiences among the young Vietnamese learners in an EFL-to-ESL transitioning context. First, while students rated AI usage and involvement to be of average or higher, their ESL readiness (including real-world language use) was moderate. This result reveals an important inconsistency between technology-mediated and authentic language learning, also found in past studies. Even when AI software allows increased access to language input and practice opportunities, it does not necessarily result in the achievement of significant communicative competence. That points to the fact that exposure alone is not enough without opportunities for purposeful, natural language use. The

second is that this study confirms that learner engagement is at the center of this relationship between AI use and ESL readiness. The association between the engagement and both readiness and real-world language use shows that the best AI tools would benefit learners if they engage them in meaningful interaction. More importantly, the findings suggest that engagement functions as a mechanism rather than merely an outcome of AI-supported learning. This finding extends existing literature by demonstrating that the effectiveness of AI is contingent upon its capacity to foster meaningful learner engagement, which subsequently facilitates ESL readiness. Consequently, engagement should be viewed as a central explanatory construct in future models of AI-assisted language learning. This complements previous studies that have shown that engaging AI can improve engagement and self-regulated learning outcomes, when learners are actively engaged in learning (Pan et al., 2024; Yang, n.d.). However, the results add to this literature, revealing that AI use is not the only factor, but engagement is a critical means by which AI affects the process of language building. Third, the qualitative investigation also highlights that students adopt AI as a cognitive support service as opposed to as communicative experience. AI helps with vocabulary, grammar, and completing tasks but can also promote passive learning behaviors, such as excessive dependence on output. This finding also raises important questions regarding the distinction between assisted performance and independent communicative competence. While AI may enhance task completion efficiency, excessive reliance on AI-generated outputs may limit opportunities for productive language use and autonomous language development. This result also parallels recent concerns expressed in the literature that AI might potentially diminish learners' active language production when it is not judiciously guided (Vistari & Yuliasri, 2026). As a result, AI may appear to offer an uneven beneficial effect; lower-level skills may only be improved, while higher-level communicative competence might not be impacted at all. A second insight is of affective nature, specifically learners' confidence and anxiety. Because there was no social pressure to practice English with AI, students felt more comfortable practicing language with AI in a positive manner. But this approach can also risk providing a framework for a lack of exposure to the randomness and complexity of real human interaction. This finding shows us that, while the artificial intelligence has the potential to simulate human communication, it is not enough to recreate the social environment of authentic linguistic production (Wang et al., 2022). Taken together, these results suggest that pedagogically structured AI-supported learning environments must support the transition to ESL. Instead of being assistive tools, AI technologies ought to be integrated into tasks that have to do with active language production, collaboration and application of language in its entirety. This change means something especially in the Vietnamese context, since in the move towards ESL the learners need not only to learn to speak and write, but also learn to use what they know in different contexts. Finally, in expanding from perceptions to intentions to readiness and language actually used, this study complements existing research, enabling a more holistic exploration of AI in language education. It also emphasizes that the success of AI is determined not only by access and frequency, but rather requires an interaction quality as well as a pedagogical design.

6. Conclusion

The study analyzed the readiness of Vietnamese university students towards English as a second language (ESL), as well as the role of AI-supported learning contexts in shaping this preparedness. The results reveal that although students often adopt AI-supported tools and show reasonable enthusiasm, their readiness for authentic language usage is at best moderate. However, such is where gaps arise between the use of digital engagement and real-world English use; this gap is especially important in the transition from EFL to ESL contexts. The

evidence indicates that AI-supported learning could have positive effects on learner engagement, self-direction, and confidence. However, the connection with communicative competence and real-life language use is less stable. The results from quantitative and qualitative research suggest that students largely utilise AI as a support device for getting through assignments, but not to develop in-depth knowledge of the language. Moreover, an over-dependence on AI and absence of legitimate encounters will prevent transfer of the learning to real-life environments. These observations highlight the role of pedagogical mediation in AI integration. AI tools on their own cannot support ESL readiness; they play a crucial role in learning in organized activities that stimulate meaningful communication and active use of language. The study contributes to the growing body of AI-assisted language learning research by proposing a process-oriented perspective in which learner engagement functions as a mediating mechanism linking AI-supported learning environments and ESL readiness. This contribution extends existing studies that have primarily focused on direct relationships between AI use and learning outcomes. Furthermore, the study provides empirical evidence from the Vietnamese higher-education context, where the transition from EFL-oriented learning toward ESL-oriented communication presents unique pedagogical challenges.

Pedagogical Implications

According to the results of this research, AI use, when applied in English language teaching should look beyond a narrow focus on convenience or efficiency to a pedagogical purposive use of AI in ways which assist the learner in developing communicative competence. In the first place, instructors must design tasks that demand engagement of language production rather than completion with AI assistance. Rather than having learners base their responses on AI-generated responses, teachers can incorporate further activities like paraphrases, oral explanations, or peer discussions that force students to make their own language through process and reproduction. This transition is needed to close the gap between aided performance and readiness for ESL. Second, the study should position AI in the role of a guide for interaction rather than an answer generation device. To facilitate AI use, teachers can frame AI experiences around interactive tasks including role-plays, simulations, or problem-solving situations for students to negotiate meaning, respond in real-time, and maintain communication. As a consequence, AI becomes a facilitator of communicative practice rather than a replacement for it. Such an approach also helps to mimic more authentic use of language, which is essential in ESL-oriented contexts. Third, the research emphasizes the need to create balanced engagement, mixing affective comfort with cognitive challenge. As AI lowers anxiety and enhances motivation, too much dependence on low-risk interaction may deprive learners of the real communicative pressure. Teachers must therefore gradually walk students through transitioning from the AI environment to natural communication environments with humans including small group discussions, presentations, and group projects to build confidence and adaptability. Another important implication relates to the emergence of AI literacy and responsible usage. Students need explicit direction on how to use AI effectively for learning rather than dependency. This involves teaching them to critically analyze AI outputs, think critically about their learning, reflect on their process, and consider the use of AI as a supplementary tool for self-regulated education, and utilize AI as a learning aid for self-regulated learning. Without that kind of guidance, AI can instead enhance passive rather than active modes of learning and may perpetuate passive learning habits. At the curriculum level, the impact of AI should be discussed within the context of ESL-oriented instructional design, which relies on the use of technology to achieve wider learning objectives, such as real-world communicative skills, intercultural competence and lifelong learning. AI should not be an

afterthought, but part of a whole-learning-process which is coherent with the transition from EFL to ESL in Vietnam. Taken together, these implications suggest that the success of AI in language learning does not merely rely on its availability but on how AI is pedagogically orchestrated for meaningful, independent, and contextually sensitive language use.

Acknowledgments

The author would like to thank the students who took part in this study voluntarily and generously shared their experiences and perspectives regarding AI-supported language learning. In addition, appreciation is extended to the colleagues and experts who contributed with valuable feedback on the research design and data collection instruments. Their assistance and constructive comments were a major factor in the completion of the study. The author is solely responsible for the content of this paper and any errors that may remain.

References

Alkandari, H. (2025). Students' learning in the time of artificial intelligence (AI): Students' perceptions of using AI tools to improve their language learning in Kuwait. *Educational Process: International Journal*. https://doi.org/10.22521/EDUPIJ.2025.15.154

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Giang, N. T., Loc, N., & Tuyen, T. (2025). Using AI-powered tools for enhancing speaking skills of EFL learners. *North American Journal of Language and Education*.

Pan, M., Guo, K., & Lai, C. (2024). Using artificial intelligence chatbots to support English-as-a-foreign-language students' self-regulated reading. *RELC Journal*. https://doi.org/10.1177/00336882241264030

Pham, L. P. C., Truong, V. L., Le, N. H. T., & Nguyen, L. (2024). The factors influencing university students' intention to adopt AI-integrated language learning methods. In *Proceedings of the 7th International Conference on Educational Technology Management*. https://doi.org/10.1145/3711403.3711444

Quang, D. N., & Minh, D. B. D. (2025). Learning English for specific purposes using artificial intelligence: University students' perspectives. *Vietnam Journal of Social Sciences*. 11. 173-190. https://doi.org/10.33100/tckhxhnv11.2.DangNgocQuang.etal

Shah, S. D., Khan, A., & Hamza, A. (2026). Exploring the role of AI-assisted language learning in enhancing second language acquisition and learner motivation. *Contemporary Journal of Educational Research*, 4(1), 527-541. https://contemporaryjournal.com/index.php/14/article/view/2032

Thanh, Q. T. (2025). The impact of artificial intelligence (AI) applications on teaching quality and learner satisfaction at foreign language centers in Ho Chi Minh City. *Journal of Knowledge Learning and Science*, ISSN: 2959-6386 (online), 4(3), 156- 164. https://doi.org/10.60087/jklst.vol4.n3.016

Vistari, L., & Yuliasri, I. (2026). From digital proficiency to communicative confidence: The mediating role of attitudes in AI-assisted language learning. *Indonesian Journal of English Education*, 14(1), 95-110. https://doi.org/10.25134/englishreview.v14i1.126

Wang, X., Pang, H., Wallace, M. P., Wang, Q., & Chen, W. (2022). Learners' perceived AI presences in AI-supported language learning: A study of AI as a humanized agent from community of inquiry. *Computer Assisted Language Learning, 35*(9), 1942–1965. https://doi.org/10.1080/09588221.2022.2056203

Yang, L. (n.d.). Empowering the autonomous learner: How AI-assisted language learning environments shape self-regulation, autonomy, and self-directed behaviors. *Language Teaching Research*, 1-23. https://doi.org/10.1177/13621688261422129

Bionote

Vu Thanh Loan is a lecturer of English at FPT University, Vietnam. She holds a Master's degree in English Linguistics and is a Microsoft Innovative Educator Expert. Her research interests include English language teaching, technology-enhanced language learning, artificial intelligence in education, and educational policy.

#64 Bridging the Gap Between Classroom Learning and Real-World Communication in ESL: Curricula and Authentic Assessments

Trang Phan Thi Quynh[1]

Abstract

English is one of the most used languages in the world; however, in ESL classes, more time is given to grammar than to communication. This has led to a situation where ESL students know how to speak and write in English grammatically but cannot apply their knowledge in real-life situations. This study was conducted to solve this problem by designing a new ESL curriculum that emphasizes communication.

A mixed-methods, quasi-experimental design was employed with two groups of university students: an experimental group receiving instruction through communicative tasks and performance-based assessment, and a control group following a traditional approach. Data were collected through pre- and post-tests, questionnaires, interviews, and classroom observations.

The results indicate that the experimental group achieved significantly greater improvement in communicative competence, particularly in fluency and interaction. Qualitative findings also reveal increased learner confidence and engagement.

The study demonstrates that in ESL education, there is a need to transform from a test-based approach to a more student-centered approach in which learning is emphasized. This can be helpful in improving student learning in English and preparing them for an international environment.

***Keywords*:** ESL curriculum design, authentic assessment, communicative competence, real-world communication

1. Introduction

Today's global world is increasingly connected through technology and other means, and English is used in education, business, science, and technology. Because people from different cultures interact more than ever, speaking English fluently and effectively is no longer just a school subject—it is a vital skill needed to succeed in college and get a job. Because of this, universities around the world focus heavily on English language programs like English as a Second Language (ESL) and English as a Foreign Language (EFL). Their goal is to prepare students to communicate in real life. However, there is still a big gap between what these courses promise to teach and what actually happens inside the classroom.

Modern language education aims to help students build real communication skills. This means being able to use English smoothly and correctly in real-life situations (Usmonova, 2024). Yet, many English classrooms today still focus too much on grammar rules and memorizing word lists. These traditional teaching methods assume that if a student understands grammar rules, they will automatically be able to speak the language well. Recent studies show that this idea is wrong. To communicate well in real life, students need active practice, natural

[1] *University of Labor and Social Affairs (Campus 2)*

conversations, and a deep understanding of context. They cannot learn these skills through passive grammar drills (Usmonova, 2024).

This problem also shows up in the way students are tested. Traditional tests often use multiple-choice questions or fill-in-the-blank grammar tasks. While these tests are easy for teachers to grade, they only measure isolated grammar knowledge. They do not show whether a student can actually use the language to communicate in real-life situations (Panayides et al., 2025). Because of this, an odd problem happens: students get high scores on grammar tests, but they feel highly anxious, lack confidence, and freeze up when they try to speak English in real life.

To fix these issues, modern language experts suggest moving toward teaching methods that focus on communication and practical tasks (Get ready to mingle!, 2026). These methods use real-world, interactive activities—like group discussions, role-plays, problem-solving tasks, and practical writing exercises. These activities force students to practice real communication in English (Usmonova, 2024). However, even when teachers agree with these ideas, they often do not use them consistently. A major reason for this failure is that schools do not match their course goals (what is taught) with their teaching methods (how it is taught) and their tests (how it is graded) (Nikula et al., 2026).

Many global studies look at communicative teaching and real-world testing as separate ideas. There is little research on how to combine them into one system within a university course. This is a huge problem in developing education systems like Vietnam, which face fixed government rules, large classes, and traditional grammar tests. Because of this, simply copying Western teaching models does not work well.

This study fills this gap by testing a unique, fully aligned model for university English courses. What makes this framework new is that it directly connects three things: course goals, daily interactive tasks, and real-world tests. By showing how to match these parts to improve fluency, accuracy, coherence, and interaction, this paper goes beyond basic theory. It provides a practical blueprint to help universities overcome local challenges and build real-world communication skills.

The study is based on a quasi-experimental method, where one group of students is exposed to communicative learning and assessment, and another group is exposed to traditional methods. The study is based on learning outcomes and learning experiences.

The research for this study is informed by the following research questions:

1. To what extent does the integration of communicative activities and authentic assessment enhance the development of communicative competence in learners compared to the traditional approach?

2. What aspects of communicative competence develop best?

3. How do learners perceive the value of communicative activities and authentic assessment in terms of engagement, confidence, and learning effectiveness?

4. What differences can be observed in terms of interaction and participation in the classroom between the two approaches?

The purpose of answering these questions is to demonstrate how better alignment between teaching and assessment can lead to more effective communication for students in the English language. It also provides some practical suggestions for designing ESL courses that can effectively prepare learners for real-life communication.

2. Literature Review

2.1. Communicative Competence in Language Education

The concept of communicative competence has been the central idea in all the recent methods of language teaching. Communicative competence, first introduced by Dell Hymes in 1972, extends the concept of grammatical competence and means the ability to communicate in appropriate social contexts. But this concept has also been further developed by Michael Canale and Merrill Swain (1980) into the following elements: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. These elements also reveal the fact that communication does not only involve the use of correct language but also the ability to make sense out of the communication.

Though this framework has been largely accepted at a theoretical level, its practical application has been mixed at best. Many researchers have noted that the move to communicative language teaching (CLT) has been largely superficial and that form-focused instruction remains the norm (Jack C. Richards, 2006). In other words, communicative competence is largely an abstract concept and not yet a "operationalized" form of instruction.

Futhermore, recent studies show that while many schools state in their course descriptions that they want to teach communication skills, their actual classrooms tell a different story. Teachers still spend most of their time lecturing and focusing strictly on grammar rules (Usmonova, 2024). Because of this, communication skills are often treated as a nice theory on paper rather than a daily practice, leaving students unprepared for the real world.

2.2. Limitations of Traditional ESL Pedagogy

The conventional approach to ESL teaching has traditionally emphasized the importance of grammar teaching and vocabulary memorizing. Though such methods can be helpful for the acquisition of language knowledge, they have been accused of failing to encourage the use of language. In this regard, it has been argued by Claire Kramsch (1993) that language use cannot be separated from its contexts, and it cannot be developed without social interactions.

In addition to this, it has been observed that conventional methods assume a linear relationship between knowledge and use. This implies that if students are aware of the grammatical rules of a language, then they will be able to use it as well. However, it has been seen through research that this is not always true, and this has been identified as one of the primary flaws of form-focused methods, as it does not allow students to deal with the unexpected nature of communication. While these methods help students remember how the language works, actual research shows they do not help students speak naturally or spontaneously. Modern research in language learning confirms that language cannot be separated from real-life situations. Language skills are social by nature, meaning students can only truly build them by talking to others and working together to communicate (Getenet et al., 2026).

2.3. Assessment Practices and Issues of Validity

Assessment is one of the most important aspects in the development of teaching and learning. In many ESL environments, the assessment procedures continue to follow the conventional modes of teaching. Discrete point tests, such as multiple-choice questions and gap-fillings, are popularly used because they are easy and reliable. Nevertheless, these tests assess knowledge in isolated areas, not communicative ability (Lyle F. Bachman, 1990).

From a validity perspective, this poses some interesting concerns. If the aim of language learning is to promote communicative competence, then surely assessment must focus on the learner's ability to communicate. It has been stated by J. Charles Alderson (2005) that "if there is a poor match between learning goals and assessment strategies, then the potential impact of the entire curriculum can be seriously damaged." In this light, perhaps assessment strategies are not only measuring what really matters but are also leading teachers to focus on test preparation rather than promoting communicative competence.

Recent research warns that when a course aims to teach communication but uses basic grammar tests to grade students, it harms the quality of the whole educational program (Nikula et al., 2026).

2.4. Authentic Learning and Authentic Assessment

In response to the limitations of these traditional methods, there is increasing interest in the development of authentic learning and authentic assessment. This approach is based on the use of tasks that are similar to real-life communication, such as problem-solving tasks and group discussions. As noted by Grant Wiggins (1998), authentic assessment is based on the learners' ability to use their knowledge rather than their ability to recall the knowledge. Real-world testing measures a student's active ability to use all their language skills to achieve a practical goal, rather than their passive ability to repeat facts from a textbook (Panayides et al., 2025).

Authentic assessment is also believed to improve learner engagement and is effective for real-life communication. However, as noted earlier, the application of authentic assessment is challenging due to various limitations and constraints. For instance, most institutions are still using traditional methods of testing due to time and resource constraints. As such, authentic assessment is used in a limited manner to improve its effectiveness.

2.5. Synthesis and Research Gap

Previous research shows that while communication is the main goal of language learning, traditional classes and grammar tests fail to achieve it. Real-world tasks and practical testing work better, but schools rarely use them consistently.

The biggest gap in current research is the lack of connection between course goals, teaching methods, and testing. While Western studies look at these parts separately, very little research shows how to combine them into one system within Asian universities. This study fills that gap by testing a fully combined system. Instead of just suggesting communication tasks, this research proves how connecting goals, lessons, and tests directly improves student skills. To show this relationship, the framework is detailed below in Figure 1.

Figure 1

Alignment of curriculum, instructional practices, and assessment for the development of communicative competence

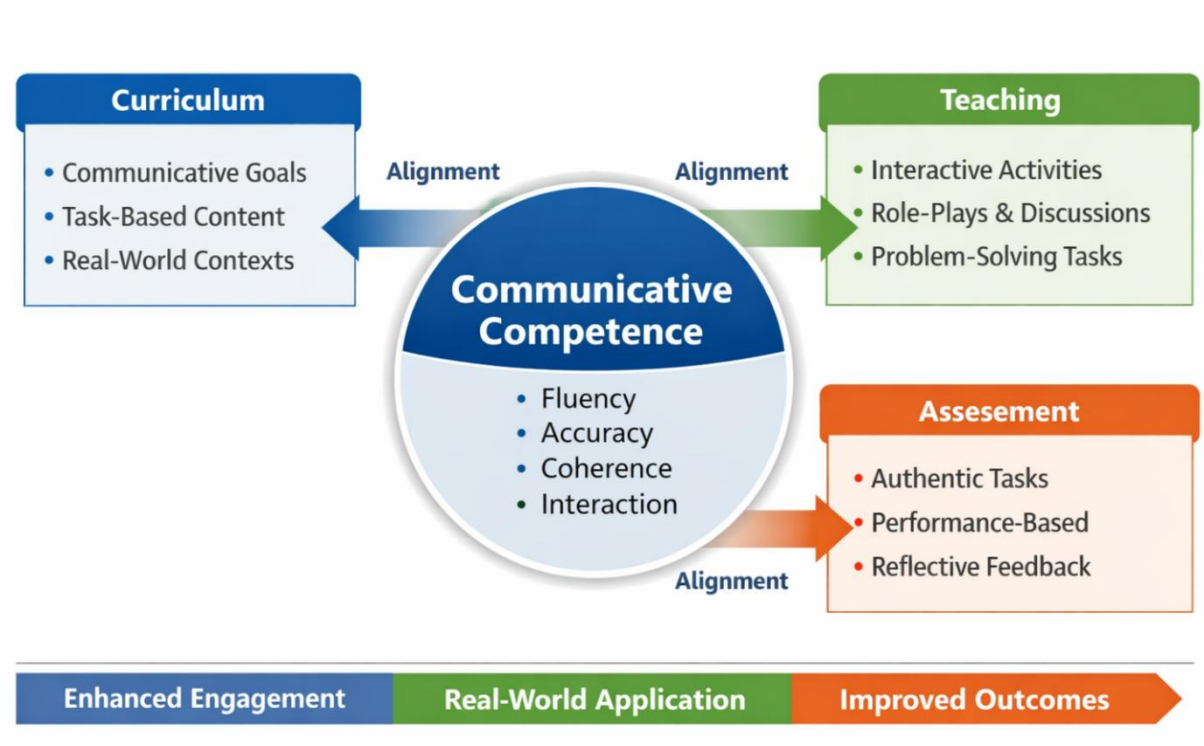


Figure 1 presents the conceptual framework that underpins this study, highlighting the need for alignment among curriculum design, instructional practices, and assessment in fostering communicative competence. At its core, the framework focuses on four key dimensions: fluency, accuracy, coherence, and interaction. It illustrates how curriculum objectives centered on real-world communication can be translated into classroom practice through communicative teaching strategies, including interactive tasks and problem-solving activities. These practices are further reinforced by authentic assessment approaches that measure learners' ability to use language in meaningful contexts. Overall, the framework suggests that learning is most effective when these elements are constructively aligned, resulting in greater student engagement and a stronger ability to apply language skills in real-life situations.

3. Methodology

3.1. Research Design

This study employs a mixed-methods research design to investigate the impact of integrating communicative activities with authentic assessment within an ESL curriculum. By combining quantitative and qualitative approaches, the study aims to provide a more comprehensive understanding of both students' learning outcomes and their classroom experiences.

More specifically, a quasi-experimental design is used, comparing a control group that receives traditional instruction with an experimental group that engages in a redesigned curriculum incorporating communicative tasks and authentic assessment practices. Quantitative data are collected to assess changes in students' communicative competence, while qualitative data offer deeper insights into learners' perspectives and the dynamics of classroom implementation. This approach is well suited to evaluating not only the effectiveness of the intervention but also how it is carried out in practice.

3.2. Participants

The study involves approximately 52 undergraduate students enrolled in an English course at a university. All participants are classified as having an intermediate level of English proficiency, based on institutional placement test results.

The students are divided into two groups. The control group (n = 26) follows a traditional instructional approach that emphasizes grammar and vocabulary, along with conventional assessment methods. In contrast, the experimental group (n = 26) is taught through a redesigned approach that incorporates communicative activities and authentic assessment tasks.

The demographic characteristics of the participants are presented in Table 1.

Table 1

Participant Demographics

Group	N	Gender (M/F)	Mean Age	English Level
Experimental	26	12 / 14	20.1	Intermediate
Control	26	11 / 15	20.3	Intermediate
Total	52	23 / 29	20.2	Intermediate

* Note. M = Male; F = Female.

All participants share comparable educational backgrounds and have studied English for a minimum of six years. Their involvement in the study is entirely voluntary, and ethical standards are carefully upheld, including the use of informed consent and the protection of participant confidentiality.

3.3. Instruments

To obtain a comprehensive set of data, the study employs multiple data collection instruments:

Pre-test and Post-test

Standardized speaking and writing assessments are administered before and after the intervention to evaluate changes in students' communicative competence. These tests are designed to reflect real-world communication tasks, such as expressing opinions, describing situations, and composing emails or brief reports.

Assessment Rubrics

Analytic rubrics are developed to assess students' performance in authentic tasks. They focus on key dimensions of communicative competence, including fluency, coherence, appropriateness, and task achievement.

Questionnaires

Structured questionnaires are distributed to students in the experimental group to explore their perceptions of the learning activities and assessment approaches. The instruments include both Likert-scale items and open-ended questions.

Classroom Observations

Classroom observations are conducted to document instructional practices, levels of student engagement, and patterns of interaction. Observation checklists are used to maintain consistency across sessions.

Interviews

Semi-structured interviews are carried out with a selected group of students and the instructor to gain deeper insights into their experiences, as well as the challenges encountered during the implementation of the redesigned curriculum.

3.4. Procedures

The study was carried out over a 10-week period and followed five main stages. Initially, a redesigned curriculum integrating communicative activities and authentic assessment was developed, and all research instruments were designed and pilot-tested. Pre-tests were then administered to both groups to determine baseline proficiency levels.

During the intervention, the control group received traditional instruction emphasizing grammar and vocabulary, whereas the experimental group participated in communicative activities, including discussions, role-plays, and problem-solving tasks, supported by authentic assessment. This phase was followed by post-testing to assess learning outcomes.

Subsequently, quantitative data from test scores were analyzed using paired and independent samples t-tests, while qualitative data from questionnaires, observations, and interviews were examined thematically to explore student engagement, perceptions, and communicative development.

4. Results

4.1. Overview of Findings

This section reports the findings of the study, with particular attention to the effects of integrating communicative activities and authentic assessment on students' communicative competence. The results are presented in two parts: (1) quantitative findings based on pre-test and post-test scores, and (2) qualitative findings derived from questionnaires, interviews, and classroom observations.

4.2. Quantitative Results

4.2.1. Pre-test and Post-test Comparison

Table 2 summarizes the mean scores and standard deviations for both the experimental and control groups in the pre-test and post-test.

Table 2

Pre-test and Post-test Scores of Experimental and Control Groups

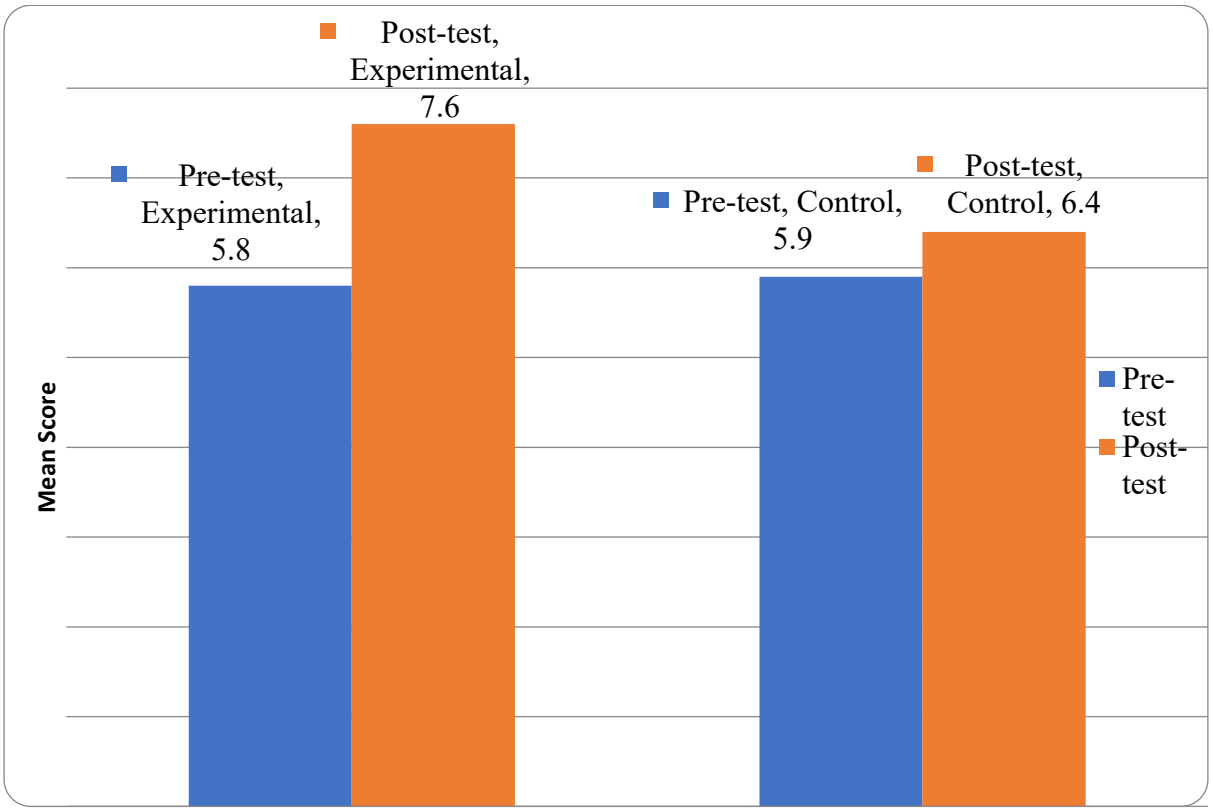
Group	Test	Mean	SD	Gain
Experimental	Pre-test	5.8	0.72	—
	Post-test	7.6	0.65	+1.8
Control	Pre-test	5.9	0.75	—
	Post-test	6.4	0.70	+0.5

* Note. Gain = difference between post-test and pre-test scores.

The differences between pre-test and post-test scores are visually illustrated in Figure 2.

Figure 2

Pre-test and post-test comparison between experimental and control groups.



The findings indicate that both groups began the study with comparable levels of communicative competence. Following the intervention, however, the experimental group showed a marked improvement, increasing from 5.8 to 7.6, whereas the control group exhibited only a slight gain, rising from 5.9 to 6.4. These results suggest that the integration of communicative activities and authentic assessment had a more pronounced positive impact on students' performance.

4.2.2. Statistical Significance

To examine whether the differences observed were meaningful rather than due to chance, both paired-sample and independent-sample t-tests were applied. The outcomes of these analyses are summarized in Table 3.

Table 3

Results of t-tests for Pre-test and Post-test Comparisons

Comparison	t-value	p-value	Interpretation
Experimental (Pre-test vs Post-test)	8.45	< .001	Highly significant
Control (Pre-test vs Post-test)	2.10	< .05	Slightly significant
Post-test (Experimental vs Control)	5.72	< .001	Highly significant

* Note. Statistical significance set at $p < .05$.

The results show that the experimental group improved to a much greater extent, with a highly significant difference ($p < .001$). In contrast, the control group made only a modest gain, although this change is still statistically significant ($p < .05$). In addition, the post-test comparison between the two groups reveals a significant difference, which supports the effectiveness of the intervention.

4.3. Performance by Communicative Criteria

To gain a clearer understanding of students' progress, their performance was assessed using specific criteria, including fluency, accuracy, coherence, and interaction. A detailed summary of results across these communicative aspects is presented in Table 4.

Table 4.

Mean Scores by Communicative Competence Criteria (Experimental Group)

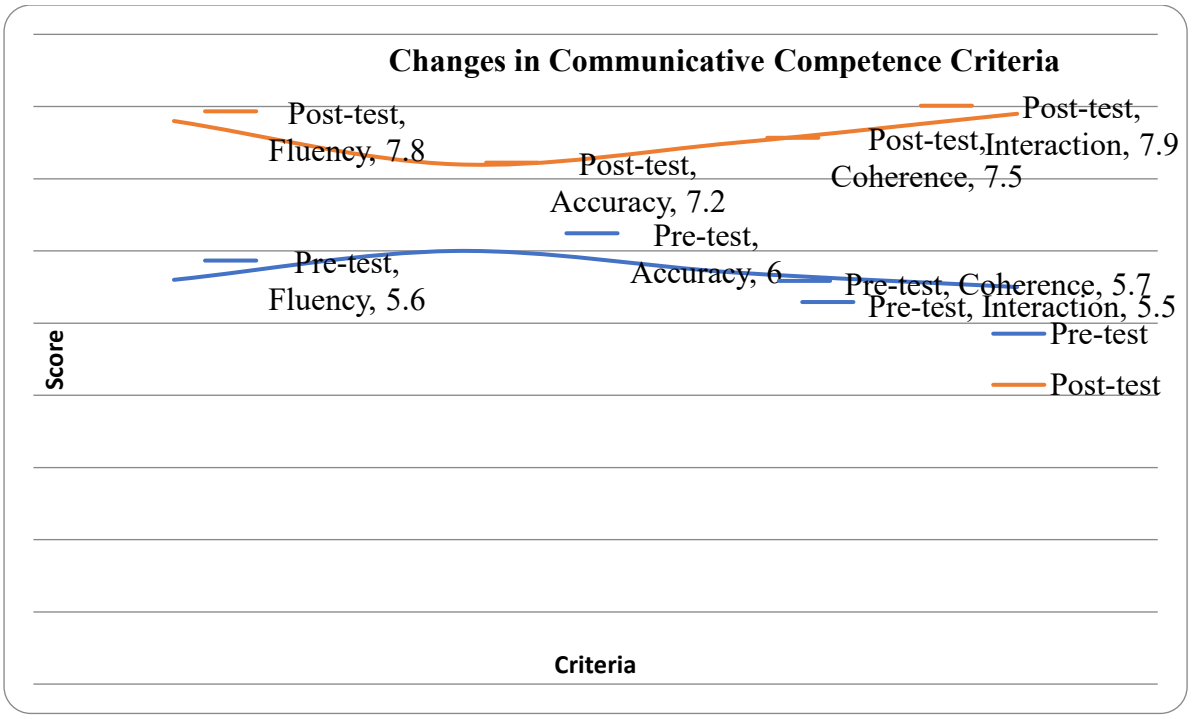
Criteria	Pre-test	Post-test	Improvement
Fluency	5.6	7.8	+2.2
Accuracy	6.0	7.2	+1.2
Coherence	5.7	7.5	+1.8
Interaction	5.5	7.9	+2.4

* Note. Scores are based on analytic rubric ratings (scale 1–10).

The changes across communicative competence criteria from pre-test to post-test are illustrated in Figure 3.

Figure 3.

Changes in communicative competence criteria from pre-test to post-test.



The most noticeable gains are in interaction and fluency, indicating that the communicative activities were especially effective in helping students participate more naturally in real-time communication. While improvements in coherence and accuracy are also present, they are slightly less pronounced. This suggests that the intervention not only strengthened students' ability to communicate but also supported their overall language development.

4.4. Qualitative Results

4.4.1. Questionnaire Findings

The questionnaire results show that most students in the experimental group had positive responses to the new approach. A summary of their perceptions is provided in Table 5.

Table 5

Student Perceptions of Communicative Tasks and Authentic Assessment

Statement	Agreement (%)
Increased confidence in speaking English	85%
Tasks are more engaging than traditional methods	82%
Assessment reflects real communicative ability	78%
Improved readiness for real-world communication	80%

* Note. Percentages represent students who selected "Agree" or "Strongly Agree."

These results suggest that authentic activities do more than improve performance; they also boost students' motivation and confidence, both of which play a crucial role in successful language learning.

4.4.2. Interview Insights

The semi-structured interviews offered a deeper understanding of students' experiences, revealing several key themes.

- Students appreciated the practical relevance of tasks (e.g., writing emails, participating in discussions).
- Many reported feeling more prepared for real-life communication.
- Some students initially found the tasks challenging but gradually became more comfortable.

These responses suggest that although authentic tasks can take some time for students to get used to, they ultimately promote meaningful learning and help develop practical language skills.

4.4.3. Classroom Observations

Classroom observations revealed notable differences between the two groups:

- The experimental group demonstrated higher levels of participation and interaction.
- Students were more willing to express ideas and engage in discussions.
- The control group showed more passive learning behaviors, focusing mainly on completing exercises.

These observations align with the quantitative results, emphasizing the important role of communicative activities in fostering active learning and increasing student engagement.

4.5. Summary of Results

Overall, the findings suggest that incorporating communicative activities and authentic assessment into the ESL curriculum leads to clear improvements in students' communicative competence. The experimental group not only achieved higher test scores than the control group but also showed greater confidence, stronger engagement, and a better ability to use language in real-life situations. These results offer strong support for the value of aligning teaching practices with authentic assessment approaches.

5. Discussion

5.1. Overview of Key Findings

The results of this study show that integrating communicative activities with authentic assessment has a clear positive effect on students' communicative competence. Compared to the control group, the experimental group made greater overall progress, especially in fluency and interaction. In addition, the qualitative findings point to higher levels of confidence, stronger engagement, and a greater sense of relevance in the learning tasks. Taken together, these results suggest that designing instruction around real-world communication needs can improve both students' language performance and their overall learning experience.

5.2. Alignment with Communicative Competence Theory

The findings of this study provide strong support for the concept of communicative competence first introduced by Dell Hymes (1972) and later expanded by Michael Canale and Merrill Swain (1980). The noticeable progress in learners' fluency, coherence, and ability to interact suggests that they were not just learning grammatical rules, but were becoming more capable of using the language in real-life situations. This reinforces the idea that communicative competence is made up of several interconnected elements and cannot be fully achieved through a sole focus on form.

In addition, the clear improvement in interactional skills underlines how important it is to give learners chances to communicate meaningfully. This aligns with the perspective of Claire Kramsch (1993), who emphasizes the social and contextual nature of language learning. Through engaging in authentic tasks, learners were able to negotiate meaning, respond to real communicative needs, and gradually build their awareness of how language is used in different contexts.

5.3. Implications for ESL Pedagogy

The findings also highlight several limitations of traditional ESL teaching methods. As suggested in earlier research, form-focused instruction often assumes that grammatical knowledge will automatically translate into communicative ability. However, the relatively modest progress observed in the control group indicates that this is not always the case in practice. This supports the view of Jack C. Richards (2006), who argues that traditional approaches alone may not be sufficient for developing real-world communication skills.

In contrast, the stronger results achieved by the experimental group point to the effectiveness of communicative language teaching (CLT), particularly when combined with authentic tasks. Engaging in activities that reflect real-life situations appears to encourage greater participation and more meaningful use of language. Interestingly, the improvement in both fluency and accuracy suggests that communicative approaches do not come at the

expense of linguistic development; instead, they seem to support it in a more integrated and balanced way.

5.4. The Role of Authentic Assessment

A key contribution of this study is its focus on authentic assessment and how it influences learning. The results make it clear that assessment is not just a way to measure outcomes—it also shapes the way students approach learning itself. As Lyle F. Bachman (1990) points out, an assessment is only truly valid if it reflects how language is used in real-life situations. In this study, performance-based tasks seemed to capture students' communicative abilities more effectively than traditional tests, which often focus more on isolated knowledge.

Students' reactions further support this point. Many of them appeared more engaged and confident when they were given opportunities to demonstrate what they could actually do with the language. This idea is consistent with Grant Wiggins (1998), who emphasizes the importance of meaningful assessment that goes beyond memorization. That said, applying this type of assessment is not without difficulties. It often demands more time, careful planning, and proper training for teachers. Because of these practical constraints, its use in everyday classrooms may still be somewhat limited, even though its benefits are quite clear.

5.5. Addressing the Curriculum–Assessment Gap

A recurring concern in the literature is the gap between what a curriculum is designed to achieve and how students are actually assessed. J. Charles Alderson (2005) highlights this issue, and the findings of this study seem to confirm it. At the same time, they also suggest a practical way to address the problem. When communicative activities were combined with authentic assessment, the revised curriculum appeared to bring learning objectives and evaluation practices into closer alignment.

Even so, alignment by itself does not guarantee success. For this approach to work effectively, it needs careful planning, clear and transparent assessment criteria, and ongoing support for teachers. Without these conditions in place, communicative methods may be applied unevenly or only at a surface level, which can reduce their overall impact.

5.6. Critical Reflection and Limitations

Although the findings are encouraging, a number of limitations need to be taken into account. To begin with, the intervention was carried out over a relatively short period of time, which makes it difficult to determine whether the observed improvements would be sustained in the long run. In addition, the study involved a fairly specific group of university students, so the results may not easily apply to other contexts or learner populations. It is also worth noting that the success of the intervention likely depended, at least in part, on the teacher's experience and the particular classroom environment, meaning that similar approaches could produce different outcomes elsewhere.

Another point to consider is that, while the study highlights the value of authentic assessment, it does not fully address some of the practical challenges teachers face. Issues such as large class sizes, institutional expectations, and the pressure of standardized testing can make it difficult to implement communicative approaches consistently. These constraints remain important obstacles and help explain why such methods are not always widely adopted in practice.

5.7. Implications for Future Research

Future research could usefully look at the longer-term effects of combining communicative activities with authentic assessment, especially across a wider range of educational settings. It would also be helpful to explore how teachers can be better supported in applying these approaches in their own classrooms, particularly in contexts where resources are limited. In addition, including learners from different backgrounds and proficiency levels could offer a broader view of how well this framework works in practice.

Taken together, this study adds to the growing body of research that supports the use of communicative teaching alongside authentic assessment in ESL education. By considering not only the strengths but also the challenges of this approach, it offers a more balanced perspective on how language programs can be designed to better equip learners for communication beyond the classroom.

6. Conclusion

6.1. Conclusion

This study set out to explore whether combining communicative activities with authentic forms of assessment could improve learners' communicative competence in ESL settings. The findings suggest that when teaching is closely connected to real-world communication, students tend to show noticeable progress, particularly in terms of fluency, interaction, and their overall ability to use the language effectively.

At the same time, the results draw attention to a gap that still exists in many ESL programs. Although communicative competence is often emphasized in curriculum goals, classroom practices do not always reflect this priority. More traditional methods of teaching and assessment remain quite common, and this can make it harder for learners to turn what they know about the language into actual communication. In contrast, the curriculum used in this study shows that when teaching and assessment are more carefully aligned, students are better able to use language in realistic situations.

The study also highlights the need to move beyond approaches that focus mainly on grammar and isolated forms. A more balanced approach—one that brings together language use, context, and interaction—seems to be more effective. Tasks that reflect real-life communication not only help students develop their skills but also appear to increase their confidence, motivation, and level of engagement. Taken together, these findings suggest that language teaching is most effective when it focuses on meaningful communication rather than treating language as a set of separate rules to be learned.

6.2. Policy Implications

The findings of this study have several implications for language education, particularly in terms of policy, curriculum design, and classroom practice.

At the curriculum level, the results point to the need for closer alignment between learning goals, teaching approaches, and assessment. It is not enough for programs to simply state that communicative competence is the objective; this aim needs to be clearly reflected in how lessons are designed and how students are evaluated. In this sense, authentic tasks should not be treated as optional add-ons, but rather as a central part of the curriculum.

From an assessment perspective, the study also raises questions about the continued reliance on traditional testing formats. While standardized tests are convenient and allow for easy comparison, they do not always capture how well learners can actually use language in

real-life situations. For this reason, there is a strong case for incorporating more performance-based and authentic forms of assessment that provide a fuller picture of students' communicative abilities.

Teacher development is another key area to consider. The success of communicative and authentic approaches depends heavily on how well they are implemented in the classroom. This means teachers need appropriate training and ongoing support, particularly in designing meaningful tasks, applying suitable assessment criteria, and managing interactive learning environments. Without this kind of preparation, even well-designed programs may fall short in practice.

At the same time, practical constraints at the institutional level cannot be overlooked. Large class sizes, limited teaching time, and restricted resources can all make it difficult to apply these approaches consistently. Institutions therefore need to create conditions that allow for flexibility and innovation, while still maintaining academic standards.

Finally, at the policy level, there is a need to move toward a more balanced view of language assessment—one that values not only grammatical accuracy but also the ability to communicate effectively. This may involve rethinking existing assessment systems so that they better reflect the kinds of language skills learners actually need outside the classroom.

6.3. Final Remarks

In closing, this study highlights the ongoing need to connect classroom learning more closely with the way language is used beyond academic settings. The results indicate that combining communicative activities with authentic assessment can have a real impact on how effectively learners develop their language abilities. Beyond its theoretical contribution, the study also offers some practical insights that may help improve current approaches to ESL teaching.

Going forward, meaningful improvement will require consistent effort across different areas, from curriculum planning to classroom implementation and broader educational policy. A more coordinated approach in these areas can better equip learners with the skills they need to communicate confidently in an increasingly interconnected world.

Acknowledgments

The author would like to express sincere gratitude to the students who participated in this study for their time, cooperation, and willingness to engage in the research activities. Their active involvement and honest feedback played an important role in shaping the findings of this study.

Special thanks are also extended to the colleagues and instructors at the University of Labor and Social Affairs (Campus 2) for their support during the implementation of the research. Their assistance in facilitating classroom activities and providing valuable insights is greatly appreciated.

The author is also grateful to those who contributed to the development of the research instruments and provided feedback throughout the research process. Their input helped improve the quality and clarity of the study.

Finally, the author would like to acknowledge the institutional support that made this research possible.

Notes

1 ESL refers to learners studying English in non-English-speaking contexts; communicative competence includes fluency, accuracy, coherence, and interaction (based on Dell Hymes; Michael Canale & Merrill Swain).

2 The study was conducted over a 10-week period with voluntary participation and ensured confidentiality.

References

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47.

Get ready to mingle! From authentic practice to authentic assessment. (2026). *Language, Culture and Curriculum*, *39*(2). https://doi.org/10.1080/17501229.2026.2660854

Getenet, S., et al. (2026). Supporting online learning in teacher education: Student-identified strategies and social constructivist implications. *Distance Education*, *47*(2). https://doi.org/10.1080/01587919.2026.2669288

Graddol, D. (2006). *English next*. British Council.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford University Press.

Nikula, T., et al. (2026). Disciplinary perspectives meet disciplinary literacies: A heuristic toolbox for content-language integration. *Journal of Multilingual and Multicultural Development*, *47*(3). https://doi.org/10.1080/01434632.2026.2646262

Panayides, P., et al. (2025). Re-introducing authentic assessment in classroom assessment courses: Finding its place in the 21st century. *Education Sciences*, *15*(11), 1564. https://doi.org/10.3390/volt222771021564

Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press.

Usmonova, M. Q. Q. (2024). Communicative competence of future teachers in the process of pedagogical education. *Journal of Academic Research in Educational Sciences*, *3*(4), 353–355.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.

Bionote

Phan Thị Quỳnh Trang is a lecturer at the University of Labor and Social Affairs (Campus 2), Vietnam. She holds a Master's degree in English Teaching Methods and has 18 years of experience in teaching English at the tertiary level. Her research interests include communicative language teaching, ESL curriculum design, and authentic assessment. She focuses on bridging the gap between classroom instruction and real-world communication, with an emphasis on improving learners' communicative competence. She has been actively involved in teaching and curriculum development and is committed to enhancing student-centered learning practices in English language education.

Appendices

Appendix A

Sample Communicative Tasks Used in the Experimental Group

Task 1: Group Discussion

Topic: The impact of social media on communication.

Students discuss in groups and respond to peers' ideas.

Task 2: Role-play Activity

Students perform real-life situations such as job interviews, customer service, or travel scenarios.

Task 3: Problem-solving Task

Students propose solutions to issues (e.g., reducing plastic waste on campus).

Task 4: Real-world Writing

Students write emails (job application, complaint response, or inquiry).

Appendix B

Analytic Assessment Rubric for Communicative Competence

Each criterion is rated on a 10-point scale. Descriptors are provided to ensure consistency in evaluation.

1. **Fluency**
   Ability to speak smoothly and maintain the flow of communication

Score Range	Descriptor
1–2	Speech is very slow, with frequent pauses; unable to maintain communication
3–4	Frequent hesitation; ideas are difficult to express continuously
5–6	Some hesitation, but able to communicate basic ideas
7–8	Generally fluent with occasional pauses; ideas expressed clearly
9–10	Speaks smoothly and naturally with minimal hesitation

2. **Accuracy**
   Correct use of grammar, vocabulary, and sentence structures

Score Range	Descriptor
1–2	Frequent errors that make understanding difficult
3–4	Many errors; meaning sometimes unclear
5–6	Some errors, but meaning is generally understandable
7–8	Minor errors that do not affect communication
9–10	Accurate use of language with very few errors

3. **Coherence**
   Clarity, organization, and logical flow of ideas

Score Range	Descriptor
1–2	Ideas are disorganized and difficult to follow
3–4	Limited organization; weak connections between ideas
5–6	Some organization; ideas are partially clear
7–8	Well-organized with clear progression of ideas
9–10	Highly coherent; ideas are logically and effectively connected

4. Interaction
Ability to engage in communication and respond appropriately

Score Range	Descriptor
1–2	Rarely responds; unable to maintain interaction
3–4	Limited participation; responses are minimal
5–6	Participates but with limited engagement
7–8	Actively participates and responds appropriately
9–10	Highly interactive; initiates and sustains communication effectively

Appendix C

Student Questionnaire on Communicative Competence and Authentic Assessment

Instructions

Please answer all questions honestly. There are no right or wrong answers. Your responses are confidential and used for research purposes only.

Part 1: Background Information (Multiple Choice)

1. Gender:

☐ Male ☐ Female ☐ Other

2. Years of learning English:

☐ Less than 5 years ☐ 5–7 years ☐ More than 7 years

3.Self-rated English level

☐ Beginner ☐ Elementary ☐ Intermediate ☐ Upper-intermediate ☐ Advanced

Part 2: Communicative Competence and Learning Experience

Scale:

1 = Strongly Disagree | 2 = Disagree | 3 = Neutral | 4 = Agree | 5 = Strongly Agree

Section A: Communicative Competence Development

Fluency

No.	Statement	1	2	3	4	5
4	I can speak English more smoothly than before.	☐	☐	☐	☐	☐
5	I can express my ideas without frequent pauses.	☐	☐	☐	☐	☐
6	I can maintain speech for a longer time without difficulty.	☐	☐	☐	☐	☐

Accuracy

No.	Statement	1	2	3	4	5
7	I use grammar more accurately when speaking English.	☐	☐	☐	☐	☐
8	I can choose appropriate vocabulary when communicating.	☐	☐	☐	☐	☐
9	My errors do not significantly affect my communication.	☐	☐	☐	☐	☐

Coherence

No.	Statement	1	2	3	4	5
10	I can organize my ideas clearly when speaking or writing.	☐	☐	☐	☐	☐

11	My ideas are logically connected when I communicate.	☐	☐	☐	☐	☐
12	I can explain my thoughts in a structured way.	☐	☐	☐	☐	☐

Interaction

No.	Statement	1	2	3	4	5
13	I can respond appropriately during conversations.	☐	☐	☐	☐	☐
14	I actively participate in discussions with others.	☐	☐	☐	☐	☐
15	I can maintain a conversation with my classmates.	☐	☐	☐	☐	☐

Section B: Confidence in Communication

No.	Statement	1	2	3	4	5
16	I feel more confident speaking English in class.	☐	☐	☐	☐	☐
17	I feel comfortable expressing my opinions in English.	☐	☐	☐	☐	☐
18	I am less afraid of making mistakes when speaking English.	☐	☐	☐	☐	☐
19	I feel prepared to use English in real-life situations.					

Section C: Authentic Assessment

No.	Statement	1	2	3	4	5
20	The assessment tasks reflect real-life communication.	☐	☐	☐	☐	☐
21	I can demonstrate my actual ability through these tasks.	☐	☐	☐	☐	☐
22	The assessment methods are fair and appropriate.	☐	☐	☐	☐	☐
23	I understand how my performance is evaluated.					

Section D: Engagement and Learning Experience

No.	Statement	1	2	3	4	5
24	The activities are more interesting than traditional lessons.	☐	☐	☐	☐	☐
25	I am motivated to participate in communicative tasks.	☐	☐	☐	☐	☐

26	I enjoy working with classmates during activities.	☐	☐	☐	☐	☐
27	The tasks encourage me to use English actively.	☐	☐	☐	☐	☐
28	I feel more engaged in this class compared to other English classes.	☐	☐	☐	☐	☐

Part 3: Learning Preferences

29. Which activity helped you most?
☐ Group discussion ☐ Role-play ☐ Problem-solving task ☐ Writing tasks ☐ Other: _____

30. Which skill improved the most?
☐ Speaking ☐ Listening ☐ Writing ☐ Vocabulary ☐ Grammar

31. What type of assessment do you prefer?
☐ Traditional tests ☐ Performance-based tasks ☐ Combination of both

Part 4: Open-ended Questions

32. Which activity helped you most in improving your communication skills? Why?
.....

33. What challenges did you face during the course?
.....

34. What suggestions do you have for improving the course?
.....

Appendix D

Participant Instructions

You will participate in classroom activities designed to improve your communication skills.

You are encouraged to actively engage in discussions, role-plays, and group tasks.

There are no “right” or “wrong” answers—focus on expressing your ideas clearly.

All responses will be kept confidential and used only for research purposes.

Appendix E

Participant Demographic Information

Total participants: 52 undergraduate students

Age range: 19–22 years

English level: Intermediate

Years of learning English: Minimum 6 years

Groups: Experimental (26), Control (26)

Appendix F

Student Questionnaire on Communicative Activities and Authentic Assessment

Instructions:

Please read each statement carefully and indicate your level of agreement. Your responses will be kept confidential and used only for research purposes.

Scale:

1 = Strongly Disagree

2 = Disagree

3 = Neutral

4 = Agree

5 = Strongly Agree

Section A: Perceived Improvement

No.	Statement	1	2	3	4	5
1	The communicative activities helped me improve my speaking skills.	☐	☐	☐	☐	☐
2	I am better able to express my ideas in English after this course.	☐	☐	☐	☐	☐
3	My fluency in English has improved.	☐	☐	☐	☐	☐
4	I can interact more confidently with others in English.	☐	☐	☐	☐	☐
5	I can organize my ideas more clearly when speaking or writing.	☐	☐	☐	☐	☐

Section B: Engagement and Motivation

No.	Statement	1	2	3	4	5
6	The activities in this course are more interesting than traditional lessons.	☐	☐	☐	☐	☐
7	I feel more motivated to participate in class.	☐	☐	☐	☐	☐
8	I enjoy working with my classmates during communicative tasks.	☐	☐	☐	☐	☐
9	The tasks encourage me to use English actively.	☐	☐	☐	☐	☐

Section C: Authentic Assessment

No.	Statement	1	2	3	4	5

10	The assessment tasks reflect real-life communication.	☐	☐	☐	☐	☐
11	I understand how my performance is evaluated.	☐	☐	☐	☐	☐
12	The assessment methods are fair.	☐	☐	☐	☐	☐
13	I can demonstrate my actual ability through these tasks.	☐	☐	☐	☐	☐

Section D: Confidence and Learning Experience

No.	Statement	1	2	3	4	5
14	I feel more confident using English in real-life situations.	☐	☐	☐	☐	☐
15	I am less afraid of making mistakes when speaking English.	☐	☐	☐	☐	☐
16	I feel more prepared for real-world communication.	☐	☐	☐	☐	☐

Section E: Open-ended Questions

17. Which activity did you find most useful? Why?

……..

18. What difficulties did you experience during the course?

……..

19. What suggestions do you have for improving the course?

……..

Appendix G

Classroom Observation Checklist

Instructions:

This checklist is used to observe classroom practices and student behaviors. Observers should tick (✓) the appropriate level and provide brief notes where necessary.

Scale

1 = Not observed 2 = Limited 3 = Moderate 4 = Frequent 5 = Highly evident

Section A: Teaching Practices

No.	Statement	1	2	3	4	5
1	Use of communicative activities (e.g., discussion, role-play)	☐	☐	☐	☐	☐
2	Use of authentic tasks (real-life scenarios)	☐	☐	☐	☐	☐
3	Teacher encourages student participation	☐	☐	☐	☐	☐
4	Teacher provides clear instructions	☐	☐	☐	☐	☐
5	Teacher facilitates interaction rather than dominates	☐	☐	☐	☐	☐

Section B: Student Engagement

No.	Statement	1	2	3	4	5
6	Students actively participate in tasks	☐	☐	☐	☐	☐
7	Students collaborate with peers	☐	☐	☐	☐	☐
8	Students show interest and motivation	☐	☐	☐	☐	☐
9	Students use English during activities	☐	☐	☐	☐	☐

Section C: Interaction Patterns

No.	Statement	1	2	3	4	5
10	Student–student interaction is evident	☐	☐	☐	☐	☐
11	Student–teacher interaction is meaningful	☐	☐	☐	☐	☐
12	Turn-taking occurs naturally	☐	☐	☐	☐	☐
13	Students respond to each other’s ideas	☐	☐	☐	☐	☐

Section D: Communicative Competence Indicators

No.	Statement	1	2	3	4	5

14	Fluency in speaking	☐	☐	☐	☐	☐
15	Accuracy of language use	☐	☐	☐	☐	☐
16	Coherence of ideas	☐	☐	☐	☐	☐
17	Ability to maintain conversation	☐	☐	☐	☐	☐

Section E: Classroom Environment

No.	Statement	1	2	3	4	5
18	Classroom atmosphere is supportive	☐	☐	☐	☐	☐
19	Students feel comfortable expressing ideas	☐	☐	☐	☐	☐
20	Errors are treated as part of learning	☐	☐	☐	☐	☐

Section F: Observer Notes

Strengths observed:

……

Challenges observed:

……

Additional comments:

……

#65 Aviation English Majors' Perceptions of the Use of AI-based Language Learning Tools for Self-regulated English Speaking Practice

Thị Nguyệt Ánh Nguyễn[1], Thị Ngọc Quý Nguyễn[2]

Abstract

Artificial Intelligence (AI) has been proved to radically enhance the quality and personalization of language education. Motivated by the growing integration of AI technologies into English learning and the need for research on learners' perceptions of their effectiveness, this study investigated how tertiary English majors perceived the application of AI-based language learning tools in English speaking self-practice. Employing a quantitative research method, the study was conducted with the participation of 161 students specializing in Aviation English at Vietnam Aviation Academy. Quantitative data on the participants' perceptions of using AI-powered software for self-regulated English speaking practice were collected using online questionnaires. The results showed that most participants strongly agreed that AI-enhanced resources provided them with personalized learning experiences and real-time feedback. However, they also identified drawbacks, including technical issues and the limited real-life applicability of examples generated by the applications. More importantly, the research detailed specific benefits and limitations of each tool investigated, as perceived by the students. These findings are expected to help any students who wish to improve their English speaking competence choose the AI software that best suits their needs and conditions to achieve their targeted goals in English learning.

***Keywords*:** AI tools, English speaking skills, language learning tools, self-regulated

1. Introduction

With the emergence of AI technologies, various English learning tools have been developed to facilitate English acquisition. These AI tools are quickly adopted by individuals and businesses worldwide for professional and personal growth (De La Vall & Araya, 2023). Given their popularity, AI tools have become one of researchers' primary concerns. Numerous research has examined the use of AI in various English learning aspects, including impacts and challenges of AI-integrated English learning (Nguyen, 2024; Quyet et al., 2024), teachers' perceptions of AI's role in language teaching (Benek, 2025; Hieu et al., 2024; Nguyen, 2024), learners' attitudes towards and experiences in using these applications (Nguyen, 2024; Benek, 2025; Nguyen & Le, 2025). Among the learners' perception-based studies, the majority were devoted to explore either learners' views on general English learning or on specific skills like writing (Thao et al., 2024). Quite few studies explored learners' perceptions of using these tools specifically for self-regulated speaking practice, however.

This study aimed to address the aforementioned research gap by exploring how Vietnam Aviation Academy (VAA) students majoring in Aviation English (AE) perceived the benefits and constraints of AI-integrated autonomous English speaking practice. It is anticipated that the research findings will provide valuable insights for AE students seeking to leverage these applications for their targeted speaking improvement.

[1] *Corresponding author, University Lecturer, Faculty of Foreign Languages, Vietnam Aviation Academy, Ho Chi Minh City, Vietnam; Email: anhntn@vaa.edu.vn*
[2] *Student, Faculty of Foreign Languages, Vietnam Aviation Academy, Ho Chi Minh City, Vietnam*

2. Literature Review

2.1. English speaking skills

Speaking skills are important in language acquisition. Effective oral communication correlates positively with academic achievement, facilitates cultural exchange and serves as a primary gateway to global knowledge (Grain et al., 2022). More importantly, good speaking proficiency is a critical determinant of career prospects, especially in aviation industry, where English serves as the standard language for customer service and passenger communication (Tangniam, 2006). Despite the critical importance of English speaking proficiency, many students continue to face significant challenges in achieving their targeted speaking goals. Determinants of suboptimal speaking performance are multifaced, encompassing not only linguistic constraints like restricted vocabulary, mispronunciation, insufficient grammar knowledge (Nhan, 2024), but also affective, environmental factors like anxiety, demotivation, fear of making mistakes (Mai et al., 2024).

Self-learning has emerged as a solution to the aforementioned constraints (Tuan, 2021). This learning mode – learners' ability to take responsibility for their learning by setting goals, selecting strategies and evaluating progress on their own (Holec, 1981) – gives learners greater control over their learning process (Buitrago, 2017). As technology develops, AI-based language learning tools become potential in providing additional ways learners can employ to improve their speaking abilities independently and efficiently (Rusmiyanto et al., 2023).

2.2. AI-based language learning tools

AI technology refers to the capability of computer systems to carry out tasks in relation to human intelligence (Russell & Norvig, 2022). The incorporation of AI in education helps create more interesting educational experiences (Lampou, 2023). AI-based language learning tools are products of such incorporation. They are computer programs, software or platforms that use AI to help with language learning (De La Vall & Araya, 2023).

The benefits of AI tools have been well documented, showing their facilitative role in English acquisition. First, AI-powered language learning software excels at personalizing learning (De La Vall & Araya, 2023). The software can identify areas requiring additional practice and improvements; then recommend customized lessons and provide exercises tailored to learners' specific needs and levels (De La Vall & Araya, 2023). Second, AI tools demonstrate significant capability in providing real-time feedback (De La Vall & Araya, 2023; Qiao & Zhao, 2023) of various types, from pronunciation, cohesion, to lexical resources and grammar (Nguyen , 2024), helping learners improve their weaknesses and progress in their language learning (De La Vall & Araya, 2023; Nguyen, 2024). Third, most AI-based applications are widely accessible, user-friendly, and inexpensive, making it possible for learners, especially those with financial problems, to easily learn anywhere at any time (De La Vall & Araya, 2023). Fourth, interactive AI features drive engagement by offering relaxing and stimulating activities (Diverkar et al., 2021), keeping learners enthusiastic and motivated (De La Vall & Araya, 2023) to maintain regular practice and complete entire courses (De La Vall & Araya, 2023). Fifth, AI tools support self-regulated learning by helping learners set goals, plan tasks, manage time and monitor progress, all of which empower them to become more independent and active in their learning (Qiao & Zhao, 2023).

However, AI-powered language learning tools are not without limitations. Firstly, AI software relies on datasets for training; consequently, their outputs – while appearing authentic – often lack originality and creativity (Poole & Polio, 2023), and can occasionally be highly superficial and formulaic (Godwin-Jones, 2024). Secondly, since AI applications have no real-

life experience, they cannot empathize, understand and use language with the same social awareness as humans; consequently, their responses, though appearing objective and accurate, are generally erroneous and inappropriate (Godwin-Jones, 2024), lack consciousness and emotion (Giannakidou & Mari, 2024), and are thus inapplicable (Godwin-Jones, 2024). Thirdly, AI tools often lack cultural sensitivity, primarily reflecting Western, male-centered values (Cao et al., 2023; Godwin-Jones, 2024), leading to significant errors in diverse cultural contexts (De La Vall & Araya, 2023). Furthermore, AI's inability to accurately replicate idiomatic and colloquial nuances frequently results in awkward or erroneous phrasing (De La Vall & Araya, 2023). Fourthly, AI tools offer limited feedback capabilities, often providing less accurate error correction and less specific explanations than human instructors (De La Vall & Araya, 2023; Godwin-Jones, 2024). Fifthly, technical barriers – such as hardware incompatibility and software instability – frequently hinder the overall effectiveness of AI in language education (Alisoy & Sadiqzade, 2024).

2.3. Previous studies on AI-based language learning tools

Recent research has extensively explored various AI tools developed for language learning, with Large Language Models (LLMs) like ChatGPT and Gemini garnering significant attention. These tools were recognized for personalizing instructions, providing real-time feedback, and fostering low-anxiety environments that built learner confidence (Abas et al., 2023; Nguyen, 2024; Celik et al., 2025; Huanca et al., 2025; Songsiengchai, 2025; Tai & Chen, 2020). Specifically, Gemini's bilingual support and authentic scenarios effectively bridged the gap between structured practice and real-world application (Huanca et al., 2025; Nguyen, 2024), while Copilot offered unique multimodal engagement through voice-to-speech processing (Kovačić & Bubaš, 2024) and Google Assistant facilitated accessible environments for autonomous (Eragamreddy, 2025; Hadi & Junor, 2022) with interactive activities (Sharizan & Yulia, 2024) and individualized practice (Fathi et al., 2025). However, LLMs faced persistent challenges, including accuracy concerns, reliability issues, and limited server capacity (Abas et al., 2023). Comparative studies further indicated that ChatGPT was inferior to specialized tools like ELSA Speak for pronunciation feedback, or Gemini in analyzing learner strengths (Vu et al., 2025).

Beyond LLMs, specialized tools such as Google Assistant, Duolingo, and ELSA Speak offered distinct benefits for language development. Google Assistant and related chatbots were praised for increasing autonomy and reducing anxiety, though they struggled with non-native accents and hardware compatibility (Setiawan & Rauf, 2025). Similarly, gamified platforms like Duolingo and ELSA Speak enhanced vocabulary, grammar, and phonological accuracy through diverse learning materials and automated feedback mechanisms (Bahtiar et al., 2025; Darsih et al., 2021; Fitria et al., 2023; Widyasari & Maghfiroh, 2023). Nevertheless, Duolingo often lacked complex, multi-turn discourse whereas ELSA Speak's speech recognition could be compromised by ambient noise in non-controlled settings (Darsih et al., 2021; Khomariyah & Ekowijayanto, 2025).

Despite the growing popularity of these AI tools, their integration into self-regulated English speaking practice remains under-researched (Tuan, 2021). Existing literature has primarily focused on general self-learning or writing improvement (Thao et al., 2024), largely neglecting learner perceptions of AI as a specific assistant for independent speaking practice. This study addressed this scholarly gap by addressing the question " What are VAA AE students' perceptions regarding the benefits and limitations of using AI-based tools within a self-regulated speaking context?"

3. Methodology

The study employed a quantitative research method to investigate how VAA AE students perceived the benefits and limitations of AI-integrated self-regulated English speaking practice since quantitative surveys are the most effective way to provide an orderly and representative picture of student attitudes (Dörnyei, 2007). Data set comprised a questionnaire administered to 161 students specializing in AE at VAA. Prior to participation, all the students were provided with a consent form, which clearly outlined the study's purpose, methodology and expected research duration. Participation was entirely voluntary.

3.1. Setting and participants

This study was conducted at VAA, targeting AE students whose future careers demand high levels of communicative competence. Despite 18 credits of the 120-credit curriculum dedicated specifically to listening and speaking skills, a significant number of students struggle to reach the required communicative proficiency levels and increasingly turn to AI tools for self-practice. However, their actual experiences remain underexplored.

Participants were recruited using a convenience sampling method (Creswell & Poth, 2017). Online consent forms were distributed to 178 students currently enrolled in the AE program at VAA, of whom 175 agreed to participate. After excluding 14 students who lacked prior experience with AI tools, the final sample consisted of 161 participants.

3.2. Instruments

Data was collected via online questionnaires to explore VAA AE students' attitudes towards and experiences with AI-integrated English speaking self-practice. The instrument was developed based on a comprehensive literature review and underwent a pilot test with three English majors who were not the main study participants to ensure its comprehensibility and validity.

The final questionnaire comprised 29 items. The initial three questions addressed the participants' background experience with AI tools. The remaining 26 items were categorized into ten thematic groups based on the literature review, namely Personalization (Q4–Q5), Immediate feedback (Q6–Q7), Accessibility (Q8–Q10), Engagement (Q11–Q12), Self-regulated learning support (Q13–Q15), Data dependency (Q16–Q17), Illusory understanding (Q18–Q20), Cultural sensitivity (Q21–Q23), Feedback accuracy (Q24–Q26), and Technical issues (Q27–Q29).

3.3. Data collection and analysis

The questionnaire was distributed online to 175 AE students who volunteered to participate in the study and 175 completed questionnaires were returned, yielding a 100% response rate. However, only data obtained from 161 of these versions were analyzed. The 14 excluded questionnaires were from students who selected "No" to the first question that asked about their prior experience with AI tools for self-regulated English speaking practice.

The data collected were descriptively analyzed and the findings were presented as percentages, using SPSS version 20.

4. Results

4.1. Use of AI-based language learning tools for English speaking self-practice

Regarding the AI tools used, ChatGPT was the most popular with 64% of the students using it, followed by Duolingo and ELSA Speak, which were used by 47.2% and 34.8% of the

participants, respectively. Gemini and Copilot, with respective usage rates of 22.4% and 18.6%, ranked fourth and fifth. Google Assistant saw lower adoption at 11.2%. Call Annie, Speechling, TalkPal, Mondly, Andy, Poe, Cici, and Coach were used by fewer than seven students each and were consequently excluded from the final analysis.

In terms of usage duration, 61.5% of participants have practiced English with AI for over six months, while only 1.9% have used these tools for less than two months. The remaining students have between two and six months of experience. This widespread, long-term engagement ensures that the participants possess sufficient experience to provide valid reflective evaluations.

4.2. Perceived benefits of AI-based language learning tools

Table 1

Students' perceptions of AI tools' benefits

Benefits	Tools	ChatGPT	Copilot	Duolingo	Elsa Speak	Gemini	Google Assistant
	N	102	30	76	56	36	18
		Agree and Strongly agree (%)					
Personalization	Q4 – identifying improvement areas	86.3	93.3	84.2	73.2	88.9	94.4
	Q5 – customizing lessons	82.4	90	84.2	82.1	86.1	94.4
Immediate feedback	Q6 – offering real-time feedback	89.2	93.3	92.1	94.6	88.9	94.4
	Q7 – pinpointing errors	89.2	96.7	82.9	82.1	80.6	94.4
Accessibility	Q8 – accessible	91.2	90	96.1	94.6	88.9	100
	Q9 – user-friendly	90.2	83.3	89.5	89.3	83.3	94.4
	Q10 - affordable	79.4	80	81.6	75	75	83.3
Engagement	Q11 – offering stimulating activities	78.4	83.3	81.6	76.8	63.9	83.3
	Q12 – encouraging regular practice	79.4	73.3	82.9	80.4	72.2	88.9
Self-regulated learning support	Q13 – setting speaking goals	83.3	86.7	86.8	80.4	80.6	94.4

	Q14 – planning learning schedules	84.3	90	81.6	76.8	77.8	94.4
	Q15 – monitoring progress	93.1	96.7	92.1	91.1	94.4	94.4

The quantitative analysis revealed that AI tools were highly regarded for their pedagogical contributions, particularly in personalizing learning experience and providing immediate feedback. Google Assistant and Copilot emerged as the most effective for diagnostic support. An overwhelming 94.4% of Google Assistant users and 93.3% of Copilot users agreed that these tools accurately identified areas for improvement. Furthermore, Copilot led the category for pinpointing mistakes of various types with a 96.7% approval rating. While ChatGPT and Duolingo maintained high consistency, averaging between 82% and 89% across feedback metrics, ELSA Speak was the top-rated for providing real-time feedback with a 94.6% approval rating. ELSA Speak was, however, perceived as less effective at broad diagnostic identification at 73.2%.

Accessibility remained a primary strength of modern AI. Google Assistant achieved a perfect score of 100% for accessibility, followed closely by Duolingo and ELSA Speak, with respective agreement rates of 96.1% and 94.6%. Despite these high scores, the data suggested a fee barrier for certain tools; Google Assistant was rated as the most affordable at 83.3%, while ELSA Speak and Gemini were perceived as slightly less accessible in terms of cost at 75%.

Engagement levels remained consistent across the board. Google Assistant and Copilot both attained an 83.3% approval rating for offering engaging, relaxing, and stimulating activities. When encouraging regular practice, Google Assistant again led with an 88.9% agreement rate, demonstrating its effectiveness in maintaining student momentum. These results suggested that the gamified elements of tools like Duolingo and the interactive nature of AI assistants were critical to keeping students motivated outside of traditional classrooms.

AI tools also served as highly effective facilitators of self-regulated learning, providing students with the structure necessary to manage their own proficiency development. Copilot demonstrated exceptional utility in this area, with 96.7% of users reporting that the tool allowed them to monitor their progress effectively. Similarly, Google Assistant achieved a 94.4% agreement rate for its capacity to assist learners in both setting personal speaking goals and planning systematic learning schedules. These tools were obviously praised for their ability to bridge the gap between classroom instruction and autonomous practice.

4.3. Perceived limitations of AI-based language learning tools

Table 2

Students' perceptions of AI tools' limitations

	Tools	ChatGPT	Copilot	Duolingo	ELSA Speak	Gemini	Google Assistant
sLimitations	N	102	30	76	56	36	18
		Agree and Strongly agree (%)					

Data dependency	Q16 – lack of originality, creativity	63.7	73.3	68.4	60.7	50	72.2
	Q17 – superficial, formulaic responses	71.6	73.3	71.1	73.2	50	61.1
Illusory understanding	Q18 – erroneous, inappropriate responses	71.6	76.7	64.5	69.6	69.4	94.4
	Q19 – unconscious, unemotional reactions	65.7	60	65.8	71.4	66.7	83.3
	Q20 – contextually inapplicable responses	52.9	53.3	57.9	46.4	44.4	72.2
Cultural sensitivity	Q21 – lack of diversity	65.7	70	67.1	62.5	61.1	72.2
	Q22 – Western, male-centered cultural bias	54.9	66.7	52.6	48.2	52.8	55.6
	Q23 - unnatural idiomatic, colloquial phrasing	66.7	66.7	63.2	69.6	72.2	83.3
Feedback accuracy	Q24 – error identification failure	54.9	63.3	59.2	51.8	50	72.2
	Q25 – error correction failure	61.8	63.3	61.8	62.5	61.1	55.6
	Q26 – unspecific explanations	66.7	70	57.9	64.3	66.7	88.9
Technical issues	Q27 – compatibility problems	49	50	48.7	60.7	55.6	44.4
	Q28 – Internet problems	63.7	66.7	64.5	53.6	63.9	50
	29 – crashes, freezing	60.8	63.3	77.6	75	55.6	83.3

Despite the pedagogical benefits of AI, participants identified several functional and systemic limitations. Across all platforms, a significant number of students reported concerns regarding data reliance and superficiality. For example, 73.3% of Copilot users felt that the speaking examples generated by the tool lacked originality and creativity, and its responses tended to be formulaic and superficial.

Furthermore, the lack of real-world or emotional understanding remained a persistent issue. Google Assistant users expressed the highest level of concern in this area, with 94.4% reporting that the tool lacked conscious and emotional reactions when conversing with human users and occasionally generated erroneous, inappropriate, and thus inapplicable responses.

Regarding cultural and linguistic sensitivity, a majority of participants across all tools indicated that AI frequently struggled with the nuances of their native language. Specifically, 83.3% of Google Assistant users and 72.2% of Gemini users noted the tools' failure to accurately phrase Vietnamese idioms and colloquial expressions. This highlighted a clear gap in the cultural adaptability of current AI models.

Finally, feedback accuracy and technical stability were identified as major barriers to long-term adoption. While tools like ChatGPT and Copilot were noted for offering overly general explanations by approximately 67% to 70% of users, technical performance also flagged. Google Assistant and Duolingo experienced the most significant reliability issues during prolonged sessions, with respective 83.3% and 77.6% of their users reporting that the applications would crash or freeze during long interaction sessions, greatly demotivating them.

5. Discussion

The analysis above showed that all the six AI tools were positively assessed by AE majors, with each tool demonstrating particular strengths in helping them self-practice English speaking skills. Google Assistant consistently outperformed other tools in learning personalization, accessibility, affordability, and motivation. These findings corroborate existing literature, which identified Google Assistant as effective in providing stress-free (Tai & Chen, 2020), accessible environments for autonomous learning (Eragamreddy, 2025; Hadi & Junor, 2022) where speaking skills were fostered (Hadi & Junor, 2022) through interactive activities (Sharizan & Yulia, 2024) and individualized practice (Fathi et al., 2025). Copilot also received positive perceptions from AE students, standing out as a premier tool for pinpointing diverse linguistic errors, offering stimulating speaking activities, and enabling independent progress monitoring. These findings highlight Copilot's significant utility for speaking practice, contrasting with previous research that primarily focused on its role in writing improvement (Abdelrady et al., 2025; Esfandiari & Akbary, 2025). ELSA Speak received positive evaluations for providing real-time feedback, although students perceived it as less effective at pinpointing specific errors. These findings partially align with previous research on the tool's automated feedback mechanisms (Bahtiar et al., 2025; Darsih et al., 2021; Widyasari & Maghfiroh, 2023).

Beyond detailing the benefits of each investigated tool, the study also identified specific limitations hindering students from fully leveraging AI tools for self-regulated speaking practice. Notably, Google Assistant, though perceived as highly effective, was identified as the most problematic. Students reported a lack of originality, frequent generation of contextually inappropriate responses, and an inability to grasp Vietnamese colloquialisms. Furthermore, the tool often provided vague feedback and suffered from recurrent technical issues. These findings reveal drawbacks beyond those previously documented by Setiawan and Rauf (2025). Copilot, though perceived as effective in specific domains, also faced criticism for producing superficial, formulaic outputs and lacking robust error-correction capabilities. Furthermore, AE students noted a Western-centric bias in vocabulary suggestions and reported that its performance was heavily dependent on stable Internet connectivity. These findings contribute significantly to the field, as Copilot remains under-researched compared to more established AI counterparts. ELSA Speak also garnered some critiques regarding software compatibility.

This finding expands upon the drawbacks identified by Darsih et al. (2021), who previously characterized ELSA Speak's primary weaknesses as an ineffective audio response system.

While early literature (i.e. Abas et al., 2023) primarily championed ChatGPT as the benchmark for AI in education, the current research findings indicate that learners are increasingly diversifying their toolkits. The superior performance of Google Assistant and Copilot in pinpointing linguistic errors suggests that as students gain experience, they prioritize diagnostic precision over the conversational breadth of general models. This shift validates the argument by Vu et al. (2025) that general LLMs often lack the specific feedback mechanisms necessary for phonological and structural refinement. The findings also demonstrate that learners are essentially curating their own AI experience, moving toward specialized assistants that function as targeted diagnostic tools rather than mere conversational partners.

Another recurring theme in the literature – most notably in Tai and Chen (2020) – is the role of AI in lowering the affective filter, or the psychological barrier that hinders language production. The current study reinforces this, as most Google Assistant and Copilot users identified the tools as providing engaging and relaxing environments for practice. By removing the fear of judgment often present in human-to-human interactions, these tools promote a sense of psychological safety.

However, the findings regarding technical friction introduce a counter-narrative to the literature. While researchers often focus on the pedagogical success of these tools, the current study highlights that infrastructure reliability remains a significant hurdle. If an AI tool crashes during prolonged practice sessions, the psychological safety it creates is immediately replaced by demotivation.

The high approval ratings for self-regulated learning – specifically the high agreement rate on progress monitoring via Copilot – strongly support the theories of Eragamreddy (2025) concerning autonomous learning. The data shows that students are not just using these tools for passive input; they are actively using them to set goals and monitor their own growth. That means AI tools empower AE students to shift from passive recipients of instructions to active managers of their own learning process. This aligns with the broader academic consensus that the most successful AI-integrated learners are those who treat the technology as a strategic assistant rather than an automated teacher (Molenaar, 2022).

6. Conclusion

6.1. Conclusion

The current study addressed the research gap concerning AI-mediated, self-regulated speaking practice among AE majors. While general research on AI in language education is increasing, limited attention has been paid to specific perceptions of students using these tools for autonomous English speaking development.

The research findings revealed that participants generally held positive perceptions of AI tools as assistants. These tools were found to enhance learning motivation and engagement by providing personalized experiences, real-time feedback, and supportive environments for self-regulation. Furthermore, accessibility and affordability were identified as key drivers of user satisfaction. Among the investigated platforms, Google Assistant and Copilot emerged as the most effective regarding pedagogical benefits, whereas ChatGPT and Duolingo remained the most popular choices despite being less prominent in terms of specific performance scores.

However, the analysis also exposed significant barriers to effective self-learning. Participants reported a heavy dependence on training data, which resulted in a perceived lack

of real understanding and cultural sensitivity. These limitations often led to unnatural or inaccurate responses. Notably, although Google Assistant and Copilot were rated as highly effective, they were also identified as the most problematic due to frequent technical issues.

Ultimately, the study suggests that while AI tools are effective in enhancing speaking competence, tool selection must be strategically aligned with individual learner's objectives and preferences.

6.2. Recommendation

Based on this study's findings, AI tools are recommended for VAA AE students seeking to self-practice their English speaking skills. However, as no single AI tool is found to be flawless, learners are encouraged to select tools that match their own learning goals, needs and preferences. For instance, Google Assistant and Copilot are ideal for AE students who prioritize immediate feedback on weaknesses, detailed suggestions for improvement or a motivating and engaging learning environment. Google Assistant is also suitable for those who prefer personalized learning experiences and accessibility. Users of Google Assistant, however, should be mindful of the potential for inaccurate feedback, unnatural responses and technical issues during practice sessions. Users of Copilot must also remain vigilant regarding the tool's tendency toward superficiality and Western-centric bias. AI-generated content should be cross-referenced with various resources to ensure cultural and technical accuracy in an aviation context.

Strategic selection of AI tools allows AE students to develop high level of English speaking proficiency to meet the demanding requirements of the aviation sector.

6.3. Limitations and directions for future research

Due to time constraints, this research exclusively analyzed quantitative data collected via online questionnaires to explore students' perceived benefits and limitations of AI tools in self-regulated English speaking practice. Follow-up interviews were not conducted, preventing a deeper investigation into underlying issues, like reasons behind the popularity of ChatGPT and Duolingo despite their less prominent benefits, or why Google Assistant and Copilot, though perceived as highly effective, are not the most popularly used. In addition, participants were recruited through convenience sampling from a single institution, which may limit the generalizability of the findings to other Aviation English students and to the broader tertiary education context in Vietnam. The specific challenges associated with the use of individual AI tools were also beyond the scope of the present study. Future research should therefore consider adopting a mixed-methods design that combines quantitative and qualitative data to provide both broader patterns and deeper insights into students' use of AI tools for self-regulated English speaking practice. In particular, qualitative interviews could help clarify the reasons behind disparities between the perceived effectiveness and popularity of specific tools, as well as explore learners' experiences, preferences, and difficulties in using individual AI tools. Expanding the sample to include students from multiple institutions would also help improve the generalizability of the findings and provide a more comprehensive understanding of AI-assisted language learning in the Vietnamese tertiary context.

References

Abas, M. A., Arumugam, S. E., Yunus, M. M., & Rafiq, K. R. M. (2023). ChatGPT and personalized learning: Opportunities and challenges in higher education. *International Journal of Academic Research in Business and Social Sciences*, *13*(12), 3536-3545.

Abdelrady, A. H., Ibrahim, D. O. O., Akram, H. (2025). Unveiling the role of Copilot in enhancing EFL learners' writing skills: A content analysis. *World Journal of English Language*, *15*(8), 174-185.

Alisoy, H., & Sadiqzade, Z. (2024). Mobile-assisted language learning (MALL): Revolutionizing language education. *Luminis Applied Science and Engineering*, *1*, 60-72.

Bahtiar, Y., Sa'diyah, L., Agustina, U. W., Nur, L. C. N., & Ma'arif, I. B. (2025). The effectiveness Of Elsa Speak application to improve speaking skill in extra class. *Pedagogi: Jurnal Ilmu Pendidikan*, *25*(1), 110-129.

Benek, K. (2025). EFL learners' and teachers' perceptions of AI-powered language learning technologies: Benefits and challenges. International *Journal of Instruction*, *18*(2), 103-120.

Buitrago, A. G. (2017). Collaborative and self-directed learning strategies to promote fluent EFL speakers. *English Language Teaching*, *10*(5), 139.

Cao, Y., Zhou, L., Lee, S., Cabello, L., Chen, M., & Hershcovich, D. (2023). Assessing cross-cultural alignment between ChatGPT and human societies: An empirical study. *arXiv preprint arXiv:2303.17466*.

Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, *8*(1), 102418-102418.

Creswell, J. W., & Poth, C.N. (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage publications.

Darsih, E., Wihadi, M., & Hanggara, A. (2021, March). Using ELSA app in speaking classes: Students' voices. In *UNISET 2020: Proceedings of the 1st Universitas Kuningan International Conference on Social Science, Environment and Technology, UNiSET 2020, 12 December 2020, Kuningan, West Java, Indonesia* (p. 295). European Alliance for Innovation.

De La Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. *The International Journal of Social Sciences and Humanities Invention*, *10*(01), 7569–7576.

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen. Y., Zhu, H.,, Hendler, J. A., & Braasch, J. (2021). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning*, *35*, 2332-2360.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Eragamreddy, N. (2025). Exploring the role of Google Assistant in enhancing EFL learners' speaking skills: A systematic review. *The JALT CALL Journal*, *21*(3), 2223-2223.

Esfandiari, R., & Akbary, O. A. (2025). Assessing the impact of Microsoft Copilot and ChatGPT on EFL learners' interactional metadiscourse in argumentative writing. *Innoeduca: International Journal of Technology and Educational Innovation*, *11*(1), 47-73.

Fathi, J., Rahimi, M., & Teo, T. (2025). Applying intelligent personal assistants to develop fluency and comprehensibility, and reduce aceentedness in EFL learners: an empirical study of Google Assistant. *Language Teaching Research*, 13621688251317786.

Fitria, N., Usman, M., & Sahara, N. (2023). Duolingo application for advancing English speaking skills. *Journal of Innovation in Educational and Cultural Research*, *4*(4), 587-595.

Giannakidou, A., & Mari, A. (2024). The human and the mechanical: logos, truthfulness, and ChatGPT. *arXiv preprint arXiv:2402.01267*.

Godwin-Jones, R. (2024). Generative AI, pragmatics, and authenticity in second language learning. arXiv preprint arXiv:2410.14395. https://doi.org/10.48550/arxiv.2410.14395.

Grain, H. M. J. S., Al-gburi, G., Suleiman, O. W., Alghazali, T., Kadhim, A. J., Hassan, A. Y., Dawood, I. I. (2022). Impact of English language proficiency, multilingualism and perceived language difficulties on international student's academic performance in Iraq. *Eurasian Journal of Applied Linguistics*. https://doi.org/10.32601/ejal.911550.

Hadi, M. S., & Junor, R. S. (2022). Speaking to devices: Can we use Google Assistant for foster students' speaking skills? *Journal of Languages and Language Teaching*, *10*(4), 570-578.

Hieu, H. H., Thao, L. T., & Thuy, P. T. (2024). Exploring the integration of artificial intelligence in Vietnamese tertiary EFL education: teacher perspectives and pedagogical challenges. *Vision: Journal for Language and Foreign Language Learning, 13*(1), 55-70. https://doi.org/10.21580/vjv13i120398.

Holec, H. (1981). *Autonomy and Foreign Language Learning*. Oxford/New York: Pergamon Press.

Huanca, Y. R. C., Farinango, M. E. G., & Pachay, J. F. Z. (2025). The use of Gemini App to enhance oral communication in EFL classroom: As pedagogical innovation.: El uso de la Aplicación Gemini para mejorar la comunicación oral en las Clases de Inglés: Como Innovación Pedagógica. *Revista Científica Multidisciplinar G-nerando*, *6*(2), ág-1867.

Khomariyah, Y., & Ekowijayanto, M. (2025). Duolingo as a digital tool for enhancing English as a Foreign Language (EFL) speaking proficiency at Zain Academy. *RIGGS: Journal of Artificial Intelligence and Digital Business*, *4*(2), 6482-6492.

Kovačić, A., & Bubaš, G. (2024, November). Students' assessment of the importance of communication skills in the asynchronous use of English as a Second Language: Learning activities with ChatGPT and Copilot to develop such skills. In *Conference Proceedings. Innovation in Language Learning 2024*.

Lampou, R. (2023). The integration of artificial intelligence in education: Opportunities and challenges. *Review of Artificial Intelligence in Education*, *4*(00), e015.

Mai, L. X., Ngoc, L. K., & Thao, L. T. (2024). Factors hindering student participation in English speaking classes: Student and lecturer perceptions. *SAGE Open*, *14*(3).

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632-645.

Nguyen, H. A. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education*, *1*(2), 18–29.

Nguyen, L. A. D., & Le, T. T. P. (2025). Exploring the effects of an AI Chatbot on emotional engagement in English speaking lessons: Insights from Call Annie. *International Journal of AI in Language Education*, *2*(2), 79-99.

Nguyen, P. D. A. (2024). Gemini Google: A potential tool for English learning. Thu Dau Mot University Journal of Science, *6*(3), https://doi.org/10.37550/tdmu.ejs/2024.03.586.

Nhan, L. K. (2024). Factors affecting English speaking proficiency among English-majored students in a Vietnamese university. *International Journal of Science and Management Studies*, *7(*3).

Poole, F. J., & Polio, C. (2023). From sci-fi to the classroom. *TASK*. Journal on Task-Based Language Teaching and Learning, *3*(2), 243-272. https://doi.org/10.1075/task.00022.poo.

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, *14*.

Quyet, C. B., Minh, N. B., & Anh, N. P. (2024). Using artificial intelligence tool in studying English skills in Vietnam – an experimental research for Vietnamese high school students. *Journal of Ecohumanism*, *3*(6), 1883–1894.

Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education*, *6*(1), 750–757.

Russell, S. J. & Norvig, P. (2022). *Artificial Intelligence: A modern approach* (4th Edition). UK: Pearson.

Setiawan, T., & Rauf, W. (2025). Google Assistant implementation for listening classes in senior high school in Barru: An exploratory case study. *Jurnal Edukasi Saintifik*, *5*(1), 27-37.

Sharizan, & Yulia, Y. (2024). The use of IPAs or intelligent personal assistants (Google Assistant, Siri, Alexa, Cortana, etc.) for learning English speaking skills.

Songsiengchai, S. (2025). Implementation of Artificial Intelligence (AI): Chat GPT for effective English language learning among Thai students in higher education. *International Journal of Education and Literacy Studies*, *13*(1), 302-312.

Tai, T. -Y., & Chen, H. H. -J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, *31*(3), 1485-1502. https://doi.org/10.1080/10494820.2020.1841801.

Tangniam, T. (2006). *An analysis of English language needs for Thai Airways ground staff*. Kasetsart University.

Thao, L. T., Hieu, H. H., & Thuy, P. T. (2024). Exploring the impacts of ChatGPT in EFL writing: Student perceptions of opportunities and challenges in Vietnamese higher education. Kognisi: Jurnal Ilmu Keguruan, *1*(2), 107-124.

Tuan, D. M. (2021). Learner autonomy in English language learning: Vietnamese EFL students' perceptions and practices. *Indonesian Journal of Applied Linguistics*, *11*(2), 307-317.

Vu, T. H. D., Le, T. P. T., Nguyen, T. H., & Tong, V. B. (2025). Comparative study of feedback on IELTS speaking performance: Pre-service teachers versus automatic speech recognition applications. *VNU Journal of Foreign Studies*, *41*(3), 26-43.

Widyasari, P., & Maghfiroh, A. (2023). The advantages of artificial intelligence ELSA speak application for speaking English learners in improving pronunciation skills. *Eltt*, *9*(1), 286-292.

Yuliawati, N. L., & Suryadi, S. (2023). Students' perceptions on the use of ELSA Speak application to improve English speaking skills. *Tell-Us Journal*, *9*(4), 1114-1137.

Bionote

Nguyễn Thị Nguyệt Ánh is a seasoned English language educator with over 30 years of teaching experience. She currently serves as a lecturer at the Vietnam Aviation Academy, where she contributes to the development of English curricula. Her primary research interests lie in the field of English language teaching, with a dedicated focus on the enhancement of language skills and pronunciation. Throughout her extensive career, she has remained committed to exploring innovative pedagogical approaches that bridge the gap between classroom theory and real-world application. This study reflects her ongoing commitment to improving communicative competence in tertiary education.

Nguyễn Thị Ngọc Quý is an undergraduate student at the Vietnam Aviation Academy, majoring in English language studies. Driven by a lifelong passion for the English language, she chose this major to deepen her linguistic expertise within a specialized technical context. Alongside her studies, she has developed a keen interest in scientific research, focusing on the intersection of language acquisition and real-life communication. This paper represents her initial foray into scholarly inquiry, reflecting her commitment to contributing to the field of Applied Linguistics. Her research aims to bridge the gap between academic theory and practical language application for future aviation professionals.

#67 LEVERAGING GENERATIVE AI TO ADDRESS VIETNAMESE EFL LEARNERS' CHALLENGES IN ENGLISH LITERATURE: A CLOSE READING AND RUBRIC-BASED ASSESSMENT FRAMEWORK FOR THIRD-YEAR UNIVERSITY STUDENTS

Nguyễn Thị Thư Hiên[1]

Abstract

This study identifies the challenges faced by third-year Vietnamese EFL learners in an English Literature course and evaluates an eight-stage ChatGPT-assisted close reading framework designed to address these challenges. Drawing on a needs analysis and a pilot study with 122 third-year English majors at HUFLIT using selected passages from Charlotte Brontë's *Jane Eyre*, the study found three major areas of difficulty: linguistic complexity, cultural unfamiliarity, and procedural uncertainty in literary analysis. These challenges appear to arise not merely from general English proficiency limitations, but from the specific linguistic, cultural, and analytical demands of literary study. The proposed framework operationalizes Vygotsky's Zone of Proximal Development by alternating AI-supported activities, such as vocabulary and contextual inquiry, literary device identification, and AI-assisted reflection, with AI-free tasks that preserve independent reading, analysis, discussion, revision, and rubric-based self-evaluation. The results indicate generally favorable learner perceptions: 88.5% of participants reported increased confidence, and mean scores on all post-pilot questionnaire items exceeded 4.00. Nevertheless, because of the exploratory nature of the research and the use of descriptive statistics as well as learners' perceptions collected at one Vietnamese university, the framework should be interpreted as an exploratory, evidence-informed pedagogical model rather than statistically confirmed evidence of improvement in learning. Replication is recommended before broader implementation in other regional EFL contexts.

Keywords: Generative AI in EFL education, literature instruction, close reading, rubric-based assessment, Vietnamese tertiary EFL

1. Introduction

1.1 Background of the Study

English literature courses are an integral component of university-level English language programmes. However, studying English literary works can be difficult for third-year students of B2 and C1 levels of English language proficiency. This difficulty is partly associated with the shift from textual comprehension to more complex forms of literary interpretation and analysis. Previous research on the challenges faced by EFL students in Vietnam has confirmed that reading and analyzing English literature texts can be problematic due to the use of sophisticated literary devices and the long nature of canonized literary texts (Nguyen, 2023).

1.2 Statement of the Problem

[1] *Ho Chi Minh City University of Foreign Languages – Information Technology; Email hienntt@huflit.edu.vn*

Several issues in learning EFL literature have been highlighted in the literature. Firstly, due to the complexity of the literary texts, language learners may experience linguistic problems in studying literary texts as a result of complex vocabulary, stylistics and syntax of these texts (Paran, 2008). Secondly, due to lack of necessary knowledge about the cultures associated with literary texts, understanding and interpreting these texts become harder for learners (Brantmeier, 2005). Thirdly, EFL learners accustomed to reading texts for information sake may face significant challenges in engaging with the aesthetic analysis of the literary texts (Rosenblatt, 1994). In sum, learners may feel uncertain about how literary analysis should be structured and evaluated, making the rubric-based assessment crucial (Panadero & Jonsson, 2013).

1.3 Research Objectives

The aim of this study is:

_ to examine challenges experienced by third-year Vietnamese EFL students while studying English literature;

_ to develop a ChatGPT-assisted framework intended to support Vietnamese EFL students in studying English literature;

_ to assess Vietnamese EFL students' attitudes toward using this framework.

1.4 Significance of the Study

It can be argued that the implementation of Generative AI, especially ChatGPT, in EFL literary analysis may help overcome these difficulties. By using Vygotsky's (1978) concept of Zone of Proximal Development as a theoretical base, some recent works in the field have suggested that Generative AI may provide scaffolding to learners by supporting their efforts during the writing process with immediate formative feedback and instructions (Xiao et al., 2025). Although Vietnamese EFL students already use AI-powered tools during their writing tasks, some learners still lack proper guidance in applying these tools successfully (Nguyen et al., 2025). Generative AI in language education is a widely researched topic, but only a few publications focus specifically on foreign language literary learning - an important area of language education that has long been discussed in the pedagogical literature (Law, 2024; Paran, 2008). Therefore, this study seeks to fill the gaps by exploring Generative AI in foreign language literary learning and suggesting classroom implications.

What distinguishes the present study from the broader body of AI-assisted EFL research is its domain-specific design. Unlike general AI-assisted EFL writing frameworks that focus on composition tasks with unrestricted AI access (e.g., Xiao et al., 2025; Teng, 2024), the present framework targets a largely neglected intersection: the close reading of canonical literary texts within a structured, alternating AI/AI-free instructional sequence. Specifically, AI interaction is constrained to three pedagogically defined steps (Steps 2, 3, and 7), directly operationalising Vygotsky's (1978) ZPD scaffolding in literary-rather than general writing-contexts. Furthermore, the five-criterion analytical writing rubric developed for this study is designed specifically for EFL literary analysis tasks, as opposed to the general writing rubrics used in most AI-EFL studies. Taken together, these features represent a domain-specific contribution that existing AI-EFL frameworks have not previously addressed empirically.

1.5 Research Questions

To reach the mentioned objectives, the research is focused on the following research questions:

1. What challenges do third-year Vietnamese EFL students experience during English literature studies?

2. How do third-year Vietnamese EFL students perceive the role of the ChatGPT-integrated framework in supporting their literary analysis tasks?

3. What are third-year Vietnamese EFL students' attitudes toward the proposed framework in terms of helpfulness, confidence, and engagement?

1.6 Scope and Limitations

The current study includes third-year undergraduate English majors at HUFLIT, where *Jane Eyre* by Charlotte Brontë served as the primary literary text for the pilot implementation. The framework developed in this study was designed to support students in close reading and rubric-based literary analysis through controlled, teacher-guided use of ChatGPT. Several boundaries should be noted. Firstly, the research involved just one institutional sample made up of 122 learners, making its applicability to any other situation outside that setting rather restricted. Secondly, only descriptive statistics were used for analysis, without the application of inferential statistical methods; therefore, the observed differences between survey mean scores should be interpreted descriptively rather than as statistical evidence of learning gains. Thirdly, learners' attitudes were collected through self-report measures, which may have influenced their responses. Fourthly, the rapid development of generative AI technologies means that the findings may require future re-examination as AI tools and classroom practices continue to evolve.

1.7 Organization of the Study

This paper is organized as follows. The review of relevant literature on EFL literary instruction, rubric-based assessment, Generative AI, and close reading will be presented in Chapter 2. In Chapter 3, the methodology of the study is outlined, covering the procedures of needs analysis and pilot experiment design. Chapter 4 discusses the findings. Chapter 5 contains conclusions of the research.

2. Literature Review

The present review addresses four interconnected themes that inform the design of the proposed framework: challenges in EFL literary instruction, close reading pedagogy in L2 contexts, rubric-based formative assessment, and the use of Generative AI in EFL education. While the pedagogical foundations of the framework draw on established scholarship in literary pedagogy and formative assessment, these are situated within an evolving body of recent empirical research on AI-mediated language learning, reflecting the dual theoretical and technological grounding of the study.

2.1 Issues of EFL Literary Instruction

Literature has long been recognised as a valuable resource in EFL education, contributing to learners' linguistic, cultural, and cognitive development (Paran, 2008). However, a persistent gap exists between the intended purposes of literary instruction and the actual experiences of L2 learners. Paran (2008), drawing on an international survey of EFL practitioners, identifies an enduring tension in foreign language education between language teaching objectives and the demands of literary study - a tension that stems partly from unexamined assumptions about learners' interpretive readiness.

Many Vietnamese students studying the English language face similar difficulties when interacting with canonical literature due to its complex nature. As noted by Lazar (1993) and

Paran (2008), literary works include archaic language, highly complex constructions, and literary stylistics, which cannot be easily mastered within communicative language teaching. Culturally, the historical and social contexts embedded in literary works - particularly those from the Victorian era - present interpretive barriers for learners whose own cultural frames of reference differ substantially from those of the texts (Lazar, 1993; Brantmeier, 2005). Analytically, the transition from information-focused reading to interpretive engagement with aesthetic and symbolic dimensions of text poses a distinct pedagogical challenge (Rosenblatt, 1994) - one documented empirically among Vietnamese final-year English majors by Nguyen (2023). These three difficulty areas directly informed the needs analysis design of the present study.

2.2 Close Textual Analysis in L2 Learning Contexts

While close reading forms an integral part of literary studies, its role in second-language instructional contexts remains relatively underexplored. In this respect, it becomes necessary to draw upon Rosenblatt's (1994) transactional theory, according to which reading can be defined as an interactive process during which cultural and linguistic backgrounds of the reader affect text comprehension. Lazar (1993) develops this position practically, proposing classroom activities that facilitate learners' access to literary and cultural meaning- several of which correspond to the AI-supported contextual inquiry incorporated in Steps 2 and 3 of the present framework. Paran (2008) further supports the integration of language development objectives within literary instruction, arguing that close attention to textual language need not be subordinated to thematic or analytical goals. Together, these perspectives provide the pedagogical rationale for a close reading sequence that interleaves AI-supported language and device analysis with independent interpretive tasks, as operationalised in the eight-step framework described in Section 3.

2.3 Rubric-based Assessment in L2 Literary Writing

A further challenge in EFL literary instruction concerns assessment transparency. Panadero and Jonsson (2013), in a systematic review of rubric use in educational settings, find that rubrics support learning by making success criteria explicit, thereby enabling learners to engage in more effective self-assessment and self-regulation. This is especially significant in literary writing tasks, where learners may perceive interpretive performance as inherently subjective and therefore difficult to evaluate against fixed criteria. Black and Wiliam (1998), in their foundational synthesis of formative assessment research, similarly demonstrate that clearly communicated performance criteria enhance learner engagement and metacognitive control - effects that extend to affective dimensions of learning. In L2 contexts, this affective dimension carries particular weight: Horwitz et al. (1986) identify fear of negative evaluation as a central component of foreign language anxiety, suggesting that explicit rubric criteria may not only clarify expectations but also reduce anxiety associated with literary performance tasks. It is on this basis that rubric-based self-assessment was incorporated as Step 8 of the present framework, using five criteria designed specifically for EFL literary analysis.

2.4 Generative AI in EFL Education

The integration of Generative AI into EFL instruction has expanded rapidly since the public release of ChatGPT, prompting a growing body of empirical and review-based research directly relevant to the design of the present framework. In their scoping review of ChatGPT use in EFL writing instruction, Xiao et al. (2025) identify scaffolding within Vygotsky’s (1978) Zone of Proximal Development as a key pedagogical mechanism through which AI can support language learners, while also noting that scaffolding should be gradually reduced to prevent

over-dependence on AI-generated output. Teng (2024) extends this analysis, acknowledging the benefits of AI-assisted feedback and idea generation while highlighting risks of academic integrity violations and passive reliance on AI responses - concerns that informed the decision to restrict AI interaction to three defined steps in the present framework.

Recent research similarly suggests that AI is most pedagogically valuable when embedded within structured instructional design and accompanied by explicit teacher guidance rather than used as a standalone tool. Kohnke et al. (2023) further argue that the ethical and effective use of AI in education depends on the development of digital competencies, including the ability to critically evaluate AI-generated outputs. Similarly, Kostka and Toncelli (2023) claim that the educational potential of using ChatGPT in English language teaching becomes possible only with clear instructional guidance. Empirical studies of AI-assisted writing further support this position: Song and Song (2023) and Yan (2023) find that ChatGPT can support writing improvement, motivation, and reflection, but only under conditions of structured implementation rather than unguided use. In the Vietnamese EFL context specifically, Nguyen et al. (2025) report that while students are already active users of AI tools, they lack the pedagogical frameworks necessary to apply them effectively in academic tasks - a finding that directly motivates the teacher-designed prompt templates used in Steps 2, 3, and 7 of the present study. Woo et al. (2025) further demonstrate that explicit instruction in prompt formulation is a necessary condition for productive AI-mediated learning, providing additional support for the controlled prompting approach adopted here. Taken together, these studies establish that the pedagogical value of AI in EFL settings is contingent on structured design, teacher guidance, and learner accountability - principles that underpin the present framework's alternating AI/AI-free architecture.

Relatedly, the capacity for critical AI literacy - defined as the selective, evaluative engagement with AI-generated output rather than passive acceptance - has emerged as an important learning objective in its own right (Long & Magerko, 2020), and one that is explicitly developed through the accept/reject reflection task in Step 7.

2.5 Research Gap

To summarize, the literature review presented above highlights several aspects of EFL literary instruction, such as the difficulties of working with literary texts, the importance of structured close reading, learning advantages associated with rubrics, and the circumstances under which AI scaffolding fosters student autonomy. However, relatively little empirical work has examined how these components can be integrated within a single pedagogical framework for Vietnamese EFL literary instruction. More precisely, the novelty of the present study lies in three interlocking design features that, to the authors' knowledge, have not been combined in prior empirical work. First, AI assistance is structurally constrained to three task-specific steps using teacher-designed prompt templates, differentiating this framework from studies in which learners engage with AI freely and without pedagogical scaffolding. Second, the analytical writing rubric is calibrated specifically for EFL literary analysis, addressing five criteria - device identification, analysis of effect, author's intention, textual engagement, and analytical language - rather than the general writing criteria commonly assessed in AI-EFL writing research. Finally, the framework is designed specifically for literary texts, where linguistic, cultural, and interpretive demands require forms of scaffolding that differ from those typically reported in AI-assisted EFL writing studies.

3. Methodology

3.1 Research Design

For the purpose of this study, an exploratory mixed-methods design was adopted, combining quantitative survey data with qualitative analysis of open-ended responses. This design is suitable for applied linguistics research because it allows numerical patterns from survey data to be interpreted alongside learners' written responses (Dörnyei, 2007). In this study, the needs analysis survey informed the framework design, while the post-pilot survey evaluated learners' perceptions after implementation.

3.2 Participants and Setting

Third-year students studying English Literature as one of their modules made up the subjects of the current study. Purposeful sampling technique was employed to sample subjects who were currently undergoing the English Literature program (Dörnyei, 2007). Based on programme requirements, participants were expected to have achieved B2–C1 levels of English proficiency. This research study was carried out in one institution to ensure consistent instructional settings throughout the investigation. The consent of the participants was sought before data collection since the researcher assured them that their participation in the study will not affect their final grades. Participation in this study was completely voluntary. There were a total of 122 third-year undergraduate students enrolled in this study through the need-analysis survey, framework piloting, and post-pilot evaluation survey.

3.3 Instruments

Two paper-based surveys were designed and used by the researcher to collect data. Paper-based surveys were selected since they facilitated easier classroom implementation and consistency.

Survey 1 (Needs Analysis Survey) was distributed prior to implementing the ChatGPT-assisted close reading instruction framework. The questions focused on gathering both quantitative and qualitative data related to linguistic, cultural, and analytical difficulties learners encountered in their literary studies, experiences in performing a close reading, and their attitude towards Generative AI technologies. Data were collected using Likert scale items and open-ended questions.

Survey 2 (Post-Pilot Feedback Survey) was conducted after implementing the instructional framework and aimed at investigating the usefulness of the framework for the learners' literary learning. The question structure of both surveys enabled comparisons of learners' experiences and perceptions prior to and after implementing the ChatGPT-assisted instruction framework.

3.4 Procedures

This study involves the following five phases of investigation. The first phase involved the administration of needs analysis survey to find out what difficulties learners encounter during their EFL literary learning. Phase 2 entailed the development of the ChatGPT-assisted instruction framework with an eight-step close reading procedure and a five-criteria analytical writing rubric based on the theoretical perspectives explained above. The current ChatGPT-assisted instruction framework comprises eight sequentially ordered steps involving alternations between AI-assisted and AI-free steps. It enables learners to perform close reading gradually, thereby implementing the theory behind Vygotsky's (1978) Zone of Proximal Development. The table below provides an overview, the primary focus and AI status of each step of the current framework.

Table 1

The Eight-Step ChatGPT-Assisted Instructional Framework Overview

Step	Activity	Focus	AI status
1	First reading	Initial comprehension; personal response	AI-free
2	Vocabulary and context inquiry	Archaic language; cultural and historical background	AI-assisted
3	Literary device identification	Identifying and labelling target devices in the passage	AI-assisted
4	Independent analytical writing	First-draft paragraph; independent literary analysis	AI-free
5	Peer discussion	Sharing and critiquing analytical responses	AI-free
6	Revised independent writing	Redrafting paragraph based on peer feedback	AI-free
7	AI-assisted reflection	Evaluating and responding to ChatGPT-generated feedback	AI-assisted
8	Final synthesis and self-evaluation	Summative paragraph with rubric-based self-assessment	AI-free

Note. Steps 2, 3, and 7 entail teacher-designed prompt templates in order to facilitate specific learning purposes through AI interaction. Remaining steps do not incorporate AI support to promote independent literary analysis. Detailed instructions for classroom implementation can be found in Appendix A. Appendix B presents the analytical writing rubric.

Phase 3 of this study involved the piloting of ChatGPT-integrated close reading framework among third-year undergraduate students undertaking an English Literature course. During the pilot process, multiple passages of Charlotte Brontë's *Jane Eyre* (1847) containing extracts from Jane's staying at Thornfield Hall and the garden scene were used. Close reading activities were carried out individually and in pairs. AI interaction was integrated in some phases using teacher-created prompts while others remained explicitly AI-free. In phase 4, a post-pilot evaluation survey was distributed to students in order to find out how useful the implemented framework was. The last phase involved data analysis and results reporting.

3.5 Data Analysis

Likert-scale data collected from the two surveys were analyzed descriptively using means, standard deviations, frequencies, and percentages. In addition, conceptually related items across the two surveys were compared descriptively to examine whether the post-pilot usefulness ratings corresponded to the difficulty areas identified in the needs analysis. These comparisons between surveys did not serve as inferential pre-test post-tests since Survey 1 measured perceived difficulty while Survey 2 assessed perceived usefulness and confidence after piloting. Open-ended answers were analyzed using thematic analysis to determine recurring patterns in relation to linguistic challenges, cultural unfamiliarity, task understanding, learner confidence, and critical assessment of feedback from the AI tool.

4. Results and Discussion

This chapter outlines the results obtained through implementation of the framework. Section 4.1 reports the needs analysis survey results addressing Research Question 1: challenges experienced by third-year Vietnamese EFL undergraduate students in studying literature. Section 4.2 discusses the findings of the post-pilot evaluation for Research Questions 2 and 3 regarding the usefulness of the ChatGPT-enabled close reading approach, the function of the rubric used, and the level of engagement among students.

4.1 Needs Analysis Findings

4.1.1 Participants' profile

Phase 1 involved a total number of 122 participants who were all third-year undergraduates taking the English literature course at HUFLIT. As can be seen from Table 2 below, the majority (n=103, 84.4%) of the learners have reached level B2 (Upper-intermediate) while others were at C1 level (n=19, 15.6%). This complies with the standards established by the English program regarding language proficiency level. When asked about their experience with literary courses, 75.4% of the learners described it as "difficult" or "very difficult" while none rated it as "easy".

Table 2

Participants' Profile (N = 122)

Characteristic	**n**	**%**
Proficiency Level		
B2 (Upper-Intermediate)	103	84.4
C1 (Advanced)	19	15.6
Overall Experience with Literature Courses		
Very difficult	36	29.5
Difficult	56	45.9
Neutral	21	17.2
Manageable	9	7.4
Easy	0	0.0

4.1.2 Self-reported difficulties: survey 1, section B

In Table 3 presented below are the means of the six Likert items used to assess learners' experience in dealing with three different types of difficulties in literary analysis of English language texts: linguistic complexity (B1, B6), lack of familiarity with cultural background (B2), and problems in performing analytic tasks (B3, B4, B5).

Table 3

Summary statistics for survey 1, section B: perceived difficulties (N = 122)

Item	**M**	**SD**	**Category**
B1. Archaic vocabulary & complex sentence structures	4.31	0.72	Linguistic
B2. Cultural references (historical events, social norms)	4.18	0.81	Cultural
B3. Identifying literary devices (e.g., metaphor, symbolism, irony)	4.09	0.88	Text Analysis
B4. Approaching close reading tasks systematically	4.23	0.76	Text Analysis
B5. Uncertainty about assessment criteria	4.05	0.91	Text Analysis
B6. Writing analytical responses in English	3.97	0.95	Linguistic

Note. Items rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). M = Mean; SD = Standard Deviation.

Given the high mean value of 4.31 and low standard deviation of SD=0.72, item B1 ("Archaic vocabulary & complex sentence structures") can be regarded as the one where students report the highest level of difficulty in comprehension of literary texts. This finding

confirms Paran's (2008) statement that literary texts can pose linguistic demands higher than those of communication. The low standard deviation suggests consistent response patterns.

Coming second was the cultural challenge (item B2): M=4.18; SD=0.81. The qualitative data collected during this study also showed that many students report experiencing problems connected with cultural differences between themselves and the cultures depicted in literary works. For example, one student found that it was hard for her to understand “why characters behave the way they do” in Victorian novels. Another student mentions that she struggles to comprehend “religious and class structures in old English society”. These results confirm the findings of Lazar (1993) and Brantmeier (2005) regarding the difficulties related to cultural issues experienced by L2 readers in literary reading activities, as well as the importance of prior knowledge/background schema to comprehension.

The mean scores for all three items in Text Analysis category (B3, B4 and B5) were high at 4.05 to 4.23. Thus, lack of clarity regarding the procedure of conducting a close reading and literary analysis can be regarded as the next major difficulty area. Given that the mean value of item B4 ("Approaching close reading tasks systematically"; M=4.23, SD=0.76) was slightly higher than the mean of item B3 ("Identifying literary devices"; M=4.09, SD=0.88), it can be concluded that in addition to difficulties in understanding literary techniques, it is hard for the students to follow a systematic procedure. This was further confirmed by results obtained in Section C of the questionnaire (item C1): only 31.7% of respondents stated that they were aware of close reading techniques. These findings are consistent with the distinction established by Rosenblatt (1994) between informative and interpretive reading, as well as Nguyen’s (2023) conclusion about the challenge caused by literary devices. Finally, the high mean of item B5 (assessment uncertainty; M=4.05, SD=0.91) suggests another potential problem - unclear assessment expectations. This idea is consistent with Panadero and Jonsson’s (2013) discussion of the importance of transparency in formative assessment criteria.

To summarize, the results of needs analysis offer compelling evidence for three main difficulty areas discussed in literature - linguistic, cultural, and procedural. The last is specifically connected with text analysis techniques.

4.1.3 Prior AI use and attitudes toward Generative AI

The results gathered in Section D of Survey 1 indicate that the respondents' level of prior experience with artificial intelligence varies. Most participants (84.4%) have tried to use various AI products including ChatGPT at least once before. Out of these, 13.1% reported using AI regularly, 40.2% - sometimes, while 31.1% tried AI just once or twice. At the same time, 15.6% of the respondents indicated that they had never used any artificial intelligence applications. Among all the respondents who previously engaged in such activity (n=103), the main reasons why they employed AI were the translation of challenging parts (92.2%), searching for cultural or historical background (89.3%), developing ideas for analysis (88.3%), and understanding literary devices (86.4%) whereas receiving feedback on writing proficiency was rare (only 45.6%).

Table 4

AI tools usage experience and Generative AI attitude: survey 1, section D (N = 122)

Item / Category		
Prior AI Use (D1)	**n**	**%**
Yes, regularly	16	13.1

Yes, sometimes	49	40.2
Tried once or twice	38	31.1
Never	19	15.6
Prior Uses of AI, if applicable (D2 - select all that apply)	**n**	**%**
Translating difficult passages	95	92.2
Understanding cultural/historical context	92	89.3
Brainstorming ideas for analysis	91	88.3
Getting explanations of literary devices	89	86.4
Getting feedback on my writing	47	45.6
Attitudes Toward ChatGPT (Likert Scale, D3-D5)	**M**	**SD**
D3. ChatGPT can help me understand difficult texts	3.81	0.94
D4. I can rely on ChatGPT's explanations for academic purposes	3.12	1.03
D5. Comfortable using ChatGPT if guided by teacher	4.27	0.68

Note. Percentages for D1 are calculated from the total sample (N = 122). For D2, percentage scores are obtained from learners who had prior exposure to AI (n = 103; the 19 learners without prior experience of AI were excluded). The learners had more than one response option in their answers, and therefore, the total percentage may exceed 100%. Items in D3-D5 have been assessed based on a five-point scale (where 1 is 'strongly disagree' and 5 is 'strongly agree').

Responses for attitudes toward ChatGPT vary depending on its purpose. While item D5 (willingness to use ChatGPT if guided by the teacher) received the highest mean value (M=4.27, SD=0.68), indicating the high willingness of students to engage in AI-assisted activities under the guidance of a teacher, D4 (trusting AI-generated explanations) received the lowest score (M=3.12, SD=1.03), indicating that learners' opinions on trusting AI are quite divided. Open-ended responses given in answer to question D6 showed similar concerns about becoming overly reliant on AI technologies, as well as possible ethical issues, which is consistent with the discussion in Teng (2024) on the challenges connected with using ChatGPT in writing EFL.

4.2 Post-Pilot Evaluation Findings

4.2.1 Framework usefulness: survey 2, section E

Following the piloting of the ChatGPT-based close reading and rubric assessment, the survey was administered again to the same sample of respondents (N=122). The items and results of Section E, which evaluates the usefulness of the framework, as well as item F2 and F3 from Section F, are provided in Table 5 below. Items' mean scores were high above 4.00 (the threshold for positive evaluation) in all cases.

Table 5

Post-pilot survey 2: Framework usefulness and confidence ratings (N = 122)

Item	**M**	**SD**
Section E: Framework Usefulness		
E1. ChatGPT helped me understand vocabulary & language	4.42	0.61

Item	M	SD
Section E: Framework Usefulness		
E2. ChatGPT helped me understand cultural/historical context	4.36	0.67
E3. ChatGPT-assisted close reading improved my text analysis	4.28	0.73
E4. The rubric clearly explained assessment criteria	4.51	0.58
E5. The rubric helped me realize my strengths & weaknesses	4.33	0.64
E6. AI-integrated lessons more helpful than traditional lessons	4.17	0.82
Section F: Confidence & Peer Recommendation		
F2. More confident in close reading tasks after framework	4.38	0.69
F3. Would recommend this course format to other students	4.45	0.62

Note. Items rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). M = Mean; SD = Standard Deviation.

Among the framework-usefulness items in Section E, the highest mean score was recorded for item E4. This finding aligns with that gathered during the needs analysis, where the difficulty connected with assessing literature was also mentioned (item B5, M=4.05). The high average score of item E4 following the pilot session is consistent with Panadero & Jonsson's (2013) finding that rubrics promote self-evaluation and clarify expectations to facilitate learning, and aligns with Black & Wiliam's (1998) broader formative assessment principle that making success criteria explicit supports the development of learners' metacognitive strategies of performance control. The learners supported this view, highlighting the value of the rubric for clear expectation setting. For example, one of the respondents mentioned that it was "like a map" for him. Another participant described the rubric as a checklist. Items E1 and E2, which concern using ChatGPT for assistance in vocabulary comprehension (M=4.42) and studying the cultural background (M=4.36), were positively evaluated by the participants. Qualitative responses to these items revealed that using ChatGPT to ask follow-up questions and get explanations based on their language competence (ZPD) was also helpful for the participants. Consistent with the findings of Xiao et al. (2025) on scaffolded learning within ZPD, learners mentioned such benefits in their qualitative responses. Here are some sample responses to the open-ended question F4 ("What was the most valuable part of AI-integrated close reading for you?").

"I like the step-by-step process. It helps me understand difficult text more easy. ChatGPT also explain vocabulary and culture background clearly."

"The AI feedback was useful because I can see my mistakes quickly. It also gives me ideas when I do not know where to begin my analysis."

These responses suggest that students valued not only the step-by-step structure but also the immediate feedback provided by ChatGPT when working through literary analysis tasks.

Finally, item E3 (ChatGPT-assisted close reading improved my text analysis; M = 4.28, SD = 0.73) also received a positive rating. Since this item concerns close reading, the result suggests that learners perceived the eight-step procedure as beneficial to their analytical work. However, because the study did not include objective pre- and post-performance measures, this finding should be interpreted as perceived analytical support rather than confirmed improvement in literary competence. This interpretation is consistent with Xiao et al.'s (2025)

discussion of gradually reduced AI scaffolding as a way to support learner independence while limiting over-reliance on AI output.

A cross-survey descriptive comparison was conducted to examine whether the areas of difficulty identified in the needs analysis corresponded to the areas in which learners later perceived the framework as useful. The comparison was not intended as an inferential pre-post test, since Survey 1 measured perceived difficulty whereas Survey 2 measured perceived usefulness after the pilot. Instead, the surveys were used to explore the alignment between learners' identified needs and their perceptions of the framework's usefulness. One of the most strongly reported difficulties in Survey 1 was archaic vocabulary and complex sentence structures (B1: M = 4.31, SD = 0.72). This corresponded to a high usefulness rating for ChatGPT-based language explanations following the pilot study (E1: M = 4.42, SD = 0.61). Similarly, cultural references were identified as another major area of difficulty (B2: M = 4.18, SD = 0.81), while AI-supported explanations of cultural and historical context were also rated highly after the pilot (E2: M = 4.36, SD = 0.67). Finally, procedural difficulties in close reading were also prominent in the needs analysis (B4: M = 4.23, SD = 0.76), and ChatGPT-guided close reading received a similarly positive post-pilot rating (E3: M = 4.28, SD = 0.73). These descriptive patterns suggest that the framework was aligned with the main difficulty areas identified before implementation. However, because the paired items were designed to assess related but non-identical constructs, such comparisons should not be viewed as quantitative evidence of learning gains but rather as descriptive evidence of instructional alignment.

4.2.2 Activity preferences and confidence: survey 2, sections F and G

The results obtained in answering to item F1 of Section F are presented in Table 6 below. Respondents were required to choose the two most helpful activities from those offered within the framework (out of five options). The highest percentage of respondents chose guided close reading tasks with AI prompts (n=73, 59.8%). The second most popular task among them was conducting research on the cultural/historical context via ChatGPT (n=55; 45.1%), followed by self-assessment according to the rubric before submission (n=52; 42.6%). The task that was chosen least frequently was peer discussion following any AI-related tasks (n=22; 18.0%).

Table 6

Most Useful Framework Activities: Survey 2, Item F1 (N=122; up to 2 selections)

Activity (F1 - select up to 2)	**n**	**%**
Guided close reading tasks with AI prompts	73	59.8
Using ChatGPT to research cultural/historical context	55	45.1
Rubric-based self-assessment prior to submission	52	42.6
Using ChatGPT to explore vocabulary and language	42	34.4
Peer discussions after AI-related activities	22	18.0

Note. Percentages reflect the proportion of respondents selecting each activity and may exceed 100% as participants could choose up to two options.

The least chosen activity following an AI-related activity was that of having a discussion among peers (n=22; 18.0%). This suggests that participants did not regard peer discussion as one of the most valuable components of the framework. However, quantitative results need to be considered in connection with the results obtained through open-ended questions. Even though participants did not name discussion among peers as one of the

activities they like, some respondents noted that they enjoyed participating in such discussions, especially because the learning process included the social component. For example, one learner suggests *"maybe give more time for discussion and writing".* Another learner mentions that *"the framework should have shorter reading passages first. Some parts of Jane Eyre are still difficult and long for me to follow."* These improvement suggestions indicate that while participants valued the AI-assisted components of the framework most highly, the low selection rate for peer discussion points to a need for redesigning or better scaffolding Step 5 in future implementations to make its value more apparent to learners. Items F2 and F3 also received high mean scores, indicating generally positive learner attitudes toward the framework. Furthermore, these findings are also supported by categorical responses in Section G (post-pilot confidence in literary analysis). 88.5% of the participants said they were "much more confident" (40.2%) or "somewhat more confident" (48.4%) in approaching close reading after the application of the framework. None of them felt less confident.

Table 7

Post-Pilot Confidence in Close Reading: Survey 2, Item G1 (N=122)

Confidence Level (G1)	**n**	**%**
Much more confident	49	40.2
Somewhat more confident	59	48.4
No change	14	11.5
Less confident	0	0.0

Note. Percentages may not sum to 100 due to rounding.

From the open-ended responses to item G2, it becomes apparent that many learners consider their increased confidence in literary analysis as due to the decrease in *"reading anxiety"* and the step-by-step procedure introduced in the framework. Indeed, as one of the respondents notes, *"Having a step-by-step process helped make it (literary analysis) from overwhelming to manageable".* Responses to F6 provide evidence that in addition to overcoming reading anxiety, the students feel more confident when approaching literature classes and are more willing to discuss the text in class. For example, one respondent said that *"before (using the framework), I think literature was very difficult and stressful. But after these lessons, I feel more confident and more interested because I know how to read and analyze better now."* Another participant notes that after using the framework, he *"doesn't feel literature is only for very good students. Now I feel more comfortable asking questions and sharing my opinion about the text".*

However, 11.5% of the learners responded they experienced no change in confidence after piloting the framework. From their written feedback, this might be because some of them already possessed high levels of confidence or because the framework was tested in a relatively short timeframe.

4.2.3 Critical AI literacy and independent analysis

One interesting finding related to Step 7, during which learners critically analyzed and accepted/rejected ChatGPT suggestions based on writing justifications for their choices, is that the activity was initially difficult but later seen as positive by most participants. Out of the 122 participants who responded to the post-pilot questionnaire, 74 (60.7%) mentioned Step 7 in their free responses. Most participants noted the difficulty of the activity in its initial stages

followed by benefits they obtained from analyzing their choices of accepting or rejecting ChatGPT's suggestion. This result corroborates Nguyen et al.'s (2025) view on the importance of providing clear instruction in AI-assisted EFL studies and may contribute to alleviating over-reliance on AI-related tools observed by Teng (2024).

To further develop the analysis of the open-ended responses, Table 8 presents selected representative examples that provide insight into how the participants perceived the framework in terms of several significant aspects. These examples were chosen because they exhibit recurring patterns in the qualitative data, including linguistic and cultural support, procedural clarity, formative feedback, independent analysis, critical evaluation of AI output, and learner confidence.

Table 8

Representative Qualitative Evidence from Open-Ended Responses (Selected Examples)

Theme	**Representative open-ended response**	**Interpretation**	**Framework link**
Linguistic and cultural access	"It helps me understand difficult text more easy. ChatGPT also explain vocabulary and culture background clearly."	AI support appeared to make difficult literary language and cultural background more accessible.	Step 2
Procedural clarity	"I like the step-by-step process."	The staged sequence helped reduce uncertainty about how to approach close reading.	Steps 1-8
Formative feedback	"The AI feedback was useful because I can see my mistakes quickly."	AI feedback helped learners notice weaknesses in their analysis or writing.	Step 7
Idea generation for analysis	"It also gives me ideas when I do not know where to begin my analysis."	AI functioned as an initial scaffold for developing analytical direction, not as a replacement for final independent writing.	Step 7
Rubric clarity	"It was like a map/checklist."	The rubric made expectations visible and supported self-assessment.	Step 8
Confidence	"Having a step-by-step process helped make it (literary analysis) from overwhelming to manageable."	The framework appeared to reduce anxiety and increase confidence in literary analysis.	Steps 1-8
Cultural understanding	"I understand better why the characters acted that way after learning about Victorian society."	AI-supported contextual inquiry helped learners connect literary events with historical and cultural background knowledge.	Step 2
Independent analysis	"At first I wanted ChatGPT to give me the answer, but later I tried writing my own ideas first."	The framework encouraged learners to maintain independent analytical thinking before consulting AI support.	Steps 4-6
Critical evaluation of AI feedback	"I did not agree with all ChatGPT suggestions because some ideas did not match my interpretation of the text."	Learners demonstrated selective acceptance of AI-generated feedback rather than passive dependence.	Step 7
Learner engagement	"The activities made literature lessons more interesting because	The structured sequence appeared to increase engagement and reduce	Steps 1-8

	I could explore the text step by step."	feelings of being overwhelmed by literary texts.	

A closer reading of these qualitative excerpts helps explain how the framework addressed learners' linguistic, cultural, and procedural difficulties beyond the descriptive survey results. The response about vocabulary and cultural background suggests that Step 2 helped learners access difficult language and contextual information that might otherwise prevent meaningful engagement with the literary text. The grammatical inaccuracies in the quoted response, such as "more easy" and "ChatGPT also explain," are retained as part of the learner's original wording; they also remind readers that perceived literary engagement occurred even while English production was still developing. Procedurally, the repeated reference to the "step-by-step process" indicates that the value of the framework came not only from AI explanations but also from the sequencing of AI-assisted and AI-free tasks. Comments on AI feedback suggest that Step 7 functioned as formative support by helping learners notice mistakes and generate possible analytical directions. Importantly, responses related to independent analysis and selective acceptance of ChatGPT suggestions indicate that learners did not simply rely on AI output passively. Instead, the structured sequence appeared to guide them toward using AI as a scaffold while preserving their own interpretive judgment in the AI-free stages.

4.3 Summary

The findings show convergence between the quantitative and qualitative data across the three research questions. Needs analysis highlighted three interrelated challenges - linguistic, cultural, and procedural - which were used in designing the framework (RQ1). In relation to RQ2, post-pilot results showed that learners perceived the eight-step approach and rubric as helpful for their literary analysis assignments. Finally, high mean scores for each item assessing usefulness of the framework, along with increased confidence and lessened reading/assessment anxiety indicate positive affective responses to the framework among participants (RQ3).

5. Conclusion

5.1 Summary of Findings

This study examined the difficulties experienced by third-year Vietnamese EFL students studying English literature and developed, implemented, and evaluated a ChatGPT-integrated pedagogical framework designed to address these challenges. Based on needs analysis data collected from 122 students enrolled in HUFLIT and the implementation of the framework using excerpts from *Jane Eyre* by Charlotte Brontë, the findings address the three research questions.

First, with regard to the identification of challenges, three categories of difficulties emerged through the needs analysis, which included linguistic challenges, cultural differences, and procedural challenges in close reading and literary analysis. Such challenges align with findings from earlier research on EFL literary instruction (Paran, 2008; Lazar, 1993; Rosenblatt, 1994) and were consistently reflected in the participants' responses.

Second, as to the design of the eight-step ChatGPT-integrated pedagogical framework, the proposed framework enabled the participants to engage in literary analysis tasks by alternating between several AI-supported activities, such as vocabulary glossing, inquiry regarding cultural context, and formative feedback for literary analysis drafts, and non-AI-supported activities, in which learners performed close reading tasks independently. The designed alternation can be considered the practical manifestation of the concept of the Zone

of Proximal Development (Vygotsky, 1978). Furthermore, the rubric consisting of five criteria addressed the issue of assessment uncertainty detected in the needs analysis and demonstrated the application of Black and Wiliam's (1998) framework.

Third, in relation to affective reactions toward the pilot framework implementation, the participants displayed generally positive affective responses toward it. Most students reported increased confidence regarding close reading, with 88.5% stating that they became more confident due to the pilot implementation of the proposed framework. The responses also suggested a perceived reduction in reading and assessment anxiety among the participants. In addition, the element of artificial intelligence literacy in Step 7 was initially perceived as a difficult one; however, learners later came to evaluate it more positively, which suggests a perceived contribution to learners' critical engagement with AI feedback.

5.2 Theoretical and Pedagogical Contributions

The findings of the present study make two significant contributions. Theoretically, it applies Vygotsky's (1978) notion of the Zone of Proximal Development and Rosenblatt's (1994) transactional theory to the context of teaching EFL literature using AI technology. This study also demonstrates an application of Black and Wiliam's (1998) ideas regarding the function of transparent assessment criteria within the framework of formative assessment in EFL literary learning. Pedagogically, the study provides a concrete eight-step framework incorporating classroom procedures, AI-supported activities, teacher-designed prompt templates, and analytical rubrics, which aims to support learners in addressing the three major categories of difficulty consistently identified among the Vietnamese EFL literature learners. In contrast to many proposals for integrating AI into EFL learning, the current study's framework has been developed specifically for close reading of literary works, thus addressing the particular needs of the learners in the field.

5.3 Limitations and Directions for Future Research

This study is subject to several limitations. First, it was conducted through one pilot implementation with a single institutional sample, which limits the generalizability of the findings. Further studies incorporating several groups of participants from various institutions would provide stronger evidence regarding the educational value of the proposed framework. Second, the experiment was performed only at HUFLIT, so the results can be influenced by the institutional practices of this particular educational organization. Third, statistical analyses, such as paired-samples t-tests or Wilcoxon signed-rank tests, were not used; hence, the results cannot be regarded as statistically verified evidence of improved language skills of the students. Fourth, the experiment measured only attitudes and perceptions of the participants and not their performance based on pre- and post-pilot writing scores. Lastly, the fact that excerpts from *Jane Eyre* were used for teaching purposes may restrict the applicability of the results obtained for other novels. Further investigation into the application of the framework is needed. Replication studies across multiple Vietnamese and regional EFL contexts are needed to assess the external validity of the framework and to determine whether its classroom value extends beyond the present setting. Future studies should also retain individual-level response data and incorporate inferential statistical analyses, such as paired-samples t-tests or Wilcoxon signed-rank tests, together with objective performance measures such as pre- and post-pilot writing scores. Such designs would make it possible to examine not only learners' perceptions of usefulness, but also measurable changes in literary analysis and analytical writing performance. Further research may also investigate learners' long-term development of critical AI literacy and teachers' perceptions of implementing and sustaining the framework.

5.4 Implications

The findings of the present study have implications for three stakeholder groups in Vietnamese higher EFL education. It should be noted, however, that given the single-institution scope of the study and the absence of inferential statistical testing, the following implications should be treated as preliminary and evidence-informed rather than conclusive. First, the study provides an evidence-informed classroom framework that can be implemented easily by teachers of English literature with no need for technical background. Furthermore, the prompts and rubrics were designed for adaptation to other literary texts and contexts. Second, the findings of the needs analysis can assist curriculum designers and programme coordinators in reconsidering the structure of literature courses, incorporating close reading, AI scaffolding, and formative rubric-based assessment as structured components of literature courses. Third, this study shows policy makers in Vietnam and the wider region that AI technologies may be integrated into existing EFL programs in pedagogically structured ways that are intended to support, rather than replace, independent analytical work in literary learning; however, broader adoption should be informed by replications across diverse institutional contexts.

Since English literature courses will likely retain their place in Vietnamese higher EFL education in the future, this study suggests that Generative AI, embedded in a pedagogically structured framework, might alleviate the difficulties experienced by students during English literature study.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension. *The Modern Language Journal*, *89*(1), 37–53. https://doi.org/10.1111/j.0026-7902.2005.00264.x

Brontë, C. (1847). *Jane Eyre*. Smith, Elder & Co.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. RELC Journal, 54(2), 537–550. https://doi.org/10.1177/00336882231162868

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. TESL-EJ, 27(3), 1–19. https://doi.org/10.55593/ej.27107int

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, *6*, 100174. https://doi.org/10.1016/j.caeo.2024.100174

Lazar, G. (1993). *Literature and language teaching: A guide for teachers and trainers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511733048

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727

Nguyen Thi, X. H., Hoang Thien, H. V., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through AI-powered tools: Perceived benefits and challenges among Vietnamese EFL students. *Discover Education*, *4*, 472. https://doi.org/10.1007/s44217-025-00905-9

Nguyen, T. T. N. (2023). Exploring Vietnamese EFL students' perceptions towards literature courses in English studies program. *International Journal of Language Instruction*, *2*(3), 80–96. https://doi.org/10.54855/ijli.23234

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Paran, A. (2008). The role of literature in instructed foreign language learning and teaching: An evidence-based survey. *Language Teaching*, *41*(4), 465–496. https://doi.org/10.1017/S026144480800520X

Rosenblatt, L. M. (1994). *The reader, the text, the poem: The transactional theory of the literary work*. Southern Illinois University Press.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. Frontiers in Psychology, 14, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, *6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Woo, D. J., Guo, K., & Susanto, H. (2025). EFL secondary students' use of ChatGPT for writing task completion pathways. *The Journal of Educational Research, 118*(6), 596–609. https://doi.org/10.1080/00220671.2025.2510382

Xiao, F., Zhu, S., & Xin, W. (2025). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Review of Education*. https://doi.org/10.1177/20965311241310881

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. Education and Information Technologies, 28(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Biographical Note

Nguyen Thi Thu Hien is a lecturer at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). She earned her M.A. in Applied Linguistics from La Trobe University, Australia, in 2005. Her research interests centre on applied linguistics, TESOL, and classroom discourse. Hiên is also a translator. She has co-translated *Coaching Salespeople into Sales Champions* by Keith Rosen (Thanh Niên Publishing House, 2023). Email: hienntt@huflit.edu.vn.

QUESTIONNAIRES

SURVEY 1 - NEEDS ANALYSIS

This questionnaire is part of a research study examining Vietnamese EFL students' challenges in English Literature and their attitudes toward ChatGPT-assisted close reading. Your participation is voluntary and anonymous. The information collected will be used solely for academic research purposes. Please answer all questions honestly based on your current knowledge and experience. There are no right or wrong answers. The questionnaire should take approximately 10-15 minutes to complete.

Thank you for your valuable contribution to this research.

Section A: Learner Background

A1. How proficient are you in English at the moment (self-assessed)?

☐ B1 (intermediate)
☐ B2 (upper intermediate)
☐ C1 (advanced)
☐ C2 (proficient)

A2. How would you describe your experience with English literature courses?

☐ very difficult
☐ difficult
☐ neutral
☐ manageable
☐ easy

Section B: Difficulties with Literature

Please mark your level of agreement to the following statements (**1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree)**

Statement	**1**	**2**	**3**	**4**	**5**
B1. Archaisms and complicated structure of sentences in literary texts cause many problems for me.	☐	☐	☐	☐	☐
B2. It is hard for me to understand the cultural context of English literature (references to historical events, social practices).	☐	☐	☐	☐	☐
B3. I find it difficult to identify literary devices (e.g., metaphor, symbolism, irony) in texts.	☐	☐	☐	☐	☐
B4. I do not know how to conduct a close reading.	☐	☐	☐	☐	☐
B5. I have doubts about what is expected from me in literary analysis tasks.	☐	☐	☐	☐	☐
B6. Analytical writing in English causes some difficulties for me.	☐	☐	☐	☐	☐

B7. What makes studying English literature especially hard for you? Please explain briefly.

Answer: …….

Section C: Experience with Close Reading & Assessment

Please rate your agreement with the following statements (**1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree)**

Statement	**1**	**2**	**3**	**4**	**5**
C1. I understand the meaning of close reading and the way it should be done.	☐	☐	☐	☐	☐

C2. The rubric/assessment criteria provided by my teacher helps me understand how to improve my work.	☐	☐	☐	☐	☐
C3. I get sufficient feedback on my literary analysis tasks.	☐	☐	☐	☐	☐

C4. What kind of support or guidance would help you most in close reading tasks?
Answer: …….

Section D: Attitudes Toward Generative AI

D1. Have you used ChatGPT or other similar applications in studying English literature so far?

☐ Yes, regularly
☐ Yes, sometimes.
☐ Once or twice.
☐ Never before.

D2. If yes, for what purposes? (Select all that apply)

☐ Translating difficult passages
☐ Understanding cultural/historical context
☐ Getting explanations of literary devices
☐ Getting feedback on my writing
☐ Brainstorming analysis ideas
☐ Other: _____

Please rate your agreement with the following statements (**1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree)**

Statement	**1**	**2**	**3**	**4**	**5**
D3. I think ChatGPT can help me understand difficult literary texts.	☐	☐	☐	☐	☐
D4. I can rely on explanations and analysis provided by ChatGPT for academic purposes.	☐	☐	☐	☐	☐
D5. I would feel comfortable using ChatGPT for studying literature with my teacher's guidance.	☐	☐	☐	☐	☐

D6. What concerns, if any, do you have about using AI tools in your literature studies?
Answer: …….

SURVEY 2

POST-PILOT FEEDBACK

This questionnaire is part of a research study examining your perception of the ChatGPT-integrated close reading and rubric-based assessment framework you have just completed. Your participation is voluntary and anonymous. The information collected will be used solely for academic research purposes. Please answer all questions honestly based on your experience with the framework. There are no right or wrong answers. The questionnaire should take approximately 10-15 minutes to complete.

Thank you for your valuable contribution to this research.

Section E: Framework Usefulness

Please indicate the degree of your agreement to the following statements (**1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree)**

Statement	**1**	**2**	**3**	**4**	**5**
E1. ChatGPT integration improved my understanding of difficult vocabulary and language in literary texts.	☐	☐	☐	☐	☐
E2. ChatGPT improved my comprehension of the cultural and historical context of the texts.	☐	☐	☐	☐	☐

E3. ChatGPT-assisted close reading practice helped me enhance my abilities to analyze literary texts.	☐	☐	☐	☐	☐
E4. The rubric introduced in this course made it clearer what was expected from me in my analysis.	☐	☐	☐	☐	☐
E5. The rubric helped me realize my strengths and weaknesses in literary analysis.	☐	☐	☐	☐	☐
E6. Compared to the traditional lessons, AI-integrated lessons were more helpful for me in studying literature.	☐	☐	☐	☐	☐

Section F: Learners' Experience

F1. What activity was the most helpful for you in this course? Choose no more than two.

☐ Exploring vocabulary and language with ChatGPT
☐ Using ChatGPT to research cultural/historical context
☐ Guided close reading tasks with AI prompts
☐ Rubric-based self-assessment prior to submission
☐ Peer discussions after AI-related activities

Please indicate the degree of your agreement to the following statements (**1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree)**

Statement	**1**	**2**	**3**	**4**	**5**
F2. I felt more confident approaching close reading tasks after using the AI-integrated framework.	☐	☐	☐	☐	☐
F3. I would recommend this course format to other students studying literature.	☐	☐	☐	☐	☐

F4. What was the most valuable part of AI-integrated close reading for you?

Answer: ……

F5. What should be added or improved about the framework? Please be specific.

Answer: ……

F6. What changes occurred in your attitude to studying English and American literature after these lessons?

Answer: ……

Section G: Overall Reflection

G1. After completing these lessons, how would you rate your confidence in conducting close reading literary texts?

☐ much more confident
☐ somewhat more confident
☐ no change
☐ less confident

G2. Is there anything else you would like to share about your experience with AI-supported literature learning?

Answer: ……

Appendix A

Note: The prompt templates and samples given in this appendix are representative examples taken from just one class using the garden scene (Chapter 23). During the pilot implementation, several passages of Jane Eyre were used.

Classroom Implementation

The next section describes the steps that both the teacher and the students should take during the pilot in the EFL literature class.

Context Setting

Text: from *Jane Eyre* (during her stay at Thornfield Hall)
AI tool: ChatGPT (used only in Steps 2, 3, and 7)
Class format: individual/paired

8-Step Procedure

Step 1: First reading (individual/ no AI)

Teacher does	Students do
✓ Assigns a passage (Ch. 23-Jane and Rochester in the garden in Thornfield) ✓ Tells students: *"Read silently. Don't look anything up now."* ✓ Teacher may show an image of a Victorian era manor house	✓ Read the passage for comprehension, underlining words/sentences that puzzle ✓ Answer in 2-3 sentences: *"What do I understand so far? What puzzles me?"*

Step 2: AI-assisted vocabulary & context (individual/ AI allowed)

Teacher does	Students do
✓ Hand out a prompt template that was pre-made, such as the following: *"I am an EFL student reading Jane Eyre (1847). Please define for me the following words/phrases and provide cultural/historical background, when appropriate: (word list)"* ✓ Ensure students use ONLY the template	✓ Use prompt template to paste words to ChatGPT ✓ Note down definitions/contexts in personal vocabulary log ✓ NOTHING ELSE. No asking ChatGPT for summary or analysis now

Step 3: AI-assisted literary device identification (paired/ AI allowed)

Teacher does	Students do
✓ Hand out another prompt template, such as: "This is a passage from *Jane Eyre*. Can you tell me whether there are literary devices and, if so, briefly explain their effect? (paste passage) ✓ Conduct a short class-wide discussion, such as: "What devices did ChatGPT find?" "Do any pairs feel the AI misidentified literary devices here? Why do you say so?" "What device stands out to you most in this passage – and why?"	✓ Work in pairs to submit the passage to ChatGPT ✓ Compare the AI's findings against your first observations in Step 1 ✓ Highlight literary devices in your copy of the text

Step 4: Independent analysis (individual/ no AI)

Teacher does	Students do

✓ Collect phones/laptops or move to the 'AI free zone' ✓ Hand out writing prompt such as the following: *"Please choose ONE literary device found in the previous step. Now, in your own words, explain why Brontë used it and its effect."* (This is an important assessment point; the teacher observes student thinking.)	✓ Write a paragraph about it independently (**no AI or partner**) ✓ The purpose is to generate **learning data** about their independent capacity

Step 5: Peer discussion (paired/ small group/no AI)

Teacher does	**Students do**
✓ Conduct structured pair-share: *"Now share your paragraph with your partner. Do you agree?"*	✓ Read each other's paragraphs from Step 4 ✓ Give feedback verbally to each other - *"I think Brontë used this device because…"* - *"My partner identified…which I didn't notice…"* etc. ✓ Identify 1 thing you'd like to improve

Step 6: Revised independent writing (individual/ no AI)

Teacher does	**Students do**
✓ Allow time for students to revise their Step 4 paragraph in light of peer feedback ✓ Explain that THIS is their work now; it shows progress.	✓ Now write a new version independently ✓ **The difference** between Step 4 and Step 6 **IS** the outcome

Step 7: AI-assisted reflection & extension (individual/ AI allowed)

Teacher does	**Students do**
✓ **Hand out another template prompt:** *"This is the draft of my analysis paragraph about Jane Eyre. Can you help me improve my paragraph? Specifically, (a) suggest a way to strengthen my argument and (b) another topic I can investigate."*	✓ Paste your Step 6 paragraph to ChatGPT ✓ Get suggestions from ChatGPT and reflect in writing about whether you agree - What did the AI say? - Do I agree? - Will I use it?

Step 8: Final synthesis (individual/ no AI)

Teacher does	**Students do**
✓ Assign an end-of-unit essay/journal writing prompt (to be done independently) Example: *"Discuss how Brontë uses imagery and symbolism to reveal Jane's emotions. In particular, describe the effect of these literary devices used together on Jane's views on love and identity."*	✓ Do the assignment: • Write the final paragraph independently (no AI) • Apply the rubric to self-assess your paragraph before submission, scoring each of the five criteria and writing one sentence to justify each score • Propose one revision idea for any criterion where your score is low

Key roles of teachers during the pilot procedure

- Gatekeeper: ensures only certain steps include AI
- Prompt designer: makes sure prompts to ChatGPT are academically appropriate
- Observer: Steps 4&6 provide insight into unassisted student thinking

- Facilitator: conducts Step 5 so that students critically examine AI's output

PROMPT for step 2: List of difficult vocabulary words and phrases

I am a university EFL student in Vietnam reading *Jane Eyre* by Charlotte Brontë (published 1847). I am currently studying an excerpt describing Jane and Rochester in the garden at Thornfield Hall.

Please help me with the following words and phrases I found difficult while reading. For each item:

1. Give a simple English definition (avoid advanced vocabulary)
2. Explain the historical or cultural context from Victorian England if relevant
3. Give one short example sentence showing how the word is used

Here are my words/phrases:

e.g.: *convulsively, impetuous, vehemence, mastery, petrified, communion, reverence, anguish.*

REMEMBER:

✓ Paste only your word list into the slot at the bottom

✓ Write down the explanations in your vocabulary log by hand

✓ Ask ChatGPT again if an explanation is unclear

✗ Do NOT ask ChatGPT to summarize the plot

✗ Do NOT ask ChatGPT to analyze characters or themes

✗ Do NOT change the prompt - use the template as written

PROMPT for step 3: Identifying Literary Devices in a Passage

I am a university EFL student in Vietnam. My partner and I are doing a close reading of a passage from *Jane Eyre* by Charlotte Brontë.

Please read the passage below and help us identify literary devices. We have already studied these five devices in class: imagery, foreshadowing, symbolism, flashback, and irony.

For each device you find in the passage:

1. Name the device (only from the five listed above)
2. Quote the exact words or phrase from the passage where it appears
3. Explain in 2-3 simple sentences what effect it creates for the reader

If a device from the list does not appear in the passage, do not invent one - just skip it.

The passage:

"Jane, do you hear that nightingale singing in the wood? Listen!" In listening, I sobbed convulsively; for I could repress what I endured no longer; I was obliged to yield, and I was shaken from head to foot with acute distress. When I did speak, it was only to express an impetuous wish that I had never been born, or never come to Thornfield.

"Because you are sorry to leave it?"

The vehemence of emotion, stirred by grief and love within me, was claiming mastery, and struggling for full sway, and asserting a right to predominate, to overcome, to live, rise, and reign at last: yes, - and to speak.

"I grieve to leave Thornfield: I love Thornfield: - I love it, because I have lived in it a full and delightful life, - momentarily at least. I have not been trampled on. I have not been petrified. I have not been buried with inferior minds, and excluded from every glimpse of communion with what is bright and energetic and high. I have talked, face to face, with what I reverence, with what I delight in, - with an original, a vigorous, an expanded mind. I have known you, Mr. Rochester; and it strikes me with terror and

anguish to feel I absolutely must be torn from you for ever. I see the necessity of departure; and it is like looking on the necessity of death."

Important:

Only identify devices within this specific passage.

Do NOT give us any overall plot summary, character analysis, or thematic essay.

AFTER YOU RECEIVE THE OUTPUT, WHAT TO DO

✓ Compare AI results with your own Step 1 observations - what did you miss? What did the AI miss?

✓ Highlight and annotate your printed passage for each device found

✓ If you disagree with the AI, note your reason - your interpretation is valid

✗ Do NOT copy the AI's explanation word-for-word into your annotation

✗ Do NOT ask the AI to write a full analysis paragraph. That is Step 4, done alone.

PROMPT for step 7: Receiving Feedback on Analysis

I am a university EFL student in Vietnam. I have just written an analysis paragraph about a passage from *Jane Eyre* by Charlotte Brontë as part of a close reading assignment.

Please read my paragraph below and give me feedback in two parts only:

Part 1: Strengthen my argument

Suggest ONE specific way I can make my analysis stronger. Focus on how well I explain the effect of the literary device on the reader. Do not rewrite my paragraph for me - just give me a clear suggestion.

Part 2: Extend my thinking

Suggest ONE related theme or critical perspective from Jane Eyre that connects to what I have written. For example: gender and independence, social class, the Gothic tradition, or Jane's moral growth. Explain briefly why it is relevant to my paragraph.

Important: Do NOT rewrite my paragraph or give me a model answer. I want suggestions I can think about and decide whether to use myself.

My paragraph:

"In this passage, Brontë use imagery to show Jane's strong emotion when she have to leave Thornfield Hall. When Jane say she was "shaken from head to foot with acute distress," the reader can imagine how much pain she feel in her body. This physical imagery help us understand that Jane's sadness is not just in her mind but also in her whole body. Brontë also use the image of the nightingale singing in the wood at the beginning of the passage. This create a contrast between the beauty of nature outside and the suffering of Jane inside. Overall, the imagery in this passage make the reader feel sympathy for Jane and show that she is a woman with very deep feeling."

Another example:

"Brontë use symbolism in this passage to express Jane's feeling about her relationship with Mr. Rochester. The nightingale at the beginning of the passage is a symbol of love and sadness in Victorian literature. Rochester ask Jane to listen to the nightingale just before she tell him her true feeling, so this symbol it suggest that their love is beautiful but also painful. Another symbol is "the necessity of death" which Jane use to describe leaving Thornfield. Death here does not mean real death but it symbolise the end of something very important in Jane's life. These two symbols show that Brontë want the reader to understand that Jane's love for Rochester and for Thornfield is very deep and it cannot be easily forget."

AFTER YOU RECEIVE FEEDBACK, REFLECT IN WRITING

✓ Write: 'The AI suggested... I agree/disagree because...'

✓ Decide for yourself whether to use the suggestion in Step 8

✓ Compare AI feedback with your peer's feedback from Step 5 - any overlap?
✗ Do NOT ask the AI to rewrite your paragraph
✗ Do NOT use the theme from Part 2 without understanding why it connects

Appendix B

Analytical Writing Rubric for Close Reading Paragraph

This rubric is applied to evaluate the paragraph of analytical writing completed by students in Step 8 of the eight-step close reading process. This rubric includes five criteria that are rated using a four-level scale (0.5, 1, 1.5, and 2). The highest possible score a student can receive for this assignment is 10 points. Following formative assessment guidelines outlined in Section 2.3, this rubric makes explicit the evaluation criteria in order to facilitate self-assessment before submission.

Criterion	2 - Excellent	1.5 - Proficient	1 - Developing	0.5 - Beginning
C1. Device identification	Correct identification of the literary device; precise quotation from the text.	Correct identification; relevant quotation from the text with minor imprecision.	Literary device identified but not accurately quoted; paraphrasing occurs.	Literary device incorrect or not identified.
C2. Analysis of effect	Concise and precise description of the effects of the literary device on the reader.	Reader's effect is explained but vaguely described without specifics.	Reader's effect is briefly mentioned but not explained in detail.	No explanation of effect of device on reader; only retelling occurs.
C3. Author's intentions	Explains why the author used this device; relates to characters and themes.	Connection between the literary device and character or theme present.	Mention of the author's intentions with some imprecision (e.g. "Brontë wants to show...").	No mention of the author's intentions and the use of the device.
C4. Engagement with the text	Good engagement with specific words or phrases used by the author.	Broad engagement with quotation but focusing on its general meaning.	Presence of quotation without engagement with specific language used.	No quotation and lack of engagement with the quotation.
C5. Analytical language and coherence	Use of precise analytical language ("conveys," "implies"); logical paragraph.	For the most part coherent and clear; some imprecise language.	Frequent use of informal language, imprecise wording, and weak paragraph structure.	Paragraph lacks coherence; imprecision prevents comprehension.

Scoring guide

Each of the five criteria is assessed separately on the scale below:

Score	Level	Score	Level
2	Excellent	1	Developing
1.5	Proficient	0.5	Beginning

Grading scale (out of 10 points)

Total score	Grade
9-10	Distinction

7-8.5	Credit
5-6.5	Pass
Under 5	Needs support

Self-evaluation instructions (Step 8)
Prior to submitting your paragraph for grading, do the following:

1. Read your paragraph according to all five criteria.
2. Assess yourself in relation to each of these criteria (score of 0.5, 1, 1.5, or 2).
3. Write down one sentence explaining your score for each criterion.
4. Propose one revision idea per criterion in case of low scores.

#71 INTEGRATING CLASSDOJO INTO WARM-UP ACTIVITIES TO SUPPORT EFL LEARNERS' WRITING PERFORMANCE AT A VIETNAMESE LANGUAGE CENTER

Chi Nguyen[1], Tran Le Khanh Phuong[2]

Abstract

This study investigates the integration of ClassDojo into warm-up activities as a form of pedagogical support for EFL writing instruction in a Vietnamese language center. Rather than functioning as a direct instructional tool, ClassDojo was embedded in the pre-writing stage to enhance learners' engagement, readiness, and motivation prior to writing tasks. A quasi-experimental mixed-methods design was employed with KET-level learners at Nam Do Language Center. Data were collected through writing tests, a student questionnaire, and semi-structured interviews.

The results indicated that the experimental group demonstrated greater improvement in writing performance than the control group. This improvement was observed at the immediate post-test and was maintained at the delayed post-test, suggesting more stable short-term retention. In addition, learners reported generally positive perceptions of the ClassDojo-supported warm-up activities, particularly in terms of perceived ease of use, perceived usefulness, engagement and readiness, and motivation.

The findings suggest that ClassDojo can support writing development indirectly by enhancing the conditions under which writing occurs. The study provides practical implications for integrating simple gamified tools into the pre-writing phase to improve learner engagement and readiness in EFL writing instruction.

***Keywords**:* ClassDojo, warm-up activities, EFL writing, learner engagement, gamification

1. Introduction

Writing is widely recognized as a core skill in English as a Foreign Language (EFL) education due to its central role in academic assessment and communication(Hyland, 2003; Weigle, 2002). However, it remains one of the most challenging skills for learners, as it requires the coordination of multiple cognitive and linguistic processes, including idea generation, organization, and language accuracy (Flower & Hayes, 1981). In addition, affective factors such as anxiety and low confidence may reduce learners' willingness to engage in writing tasks (Akhtar et al., 2019; Rasool et al., 2023). As a result, writing performance is influenced not only by linguistic competence but also by learners' readiness and engagement prior to task execution.

Within this perspective, warm-up or pre-writing activities play an important role in preparing learners for writing. These activities help activate prior knowledge, recall vocabulary, and reduce anxiety before learners begin writing tasks (Yuk & Yunus, 2021).

[1]*HUFLIT University*
[2] *Graduate student, HUFLIT University*

Despite their potential, warm-up stages are often minimized in practice, particularly in time-constrained instructional settings, resulting in limited support for learners' readiness.

At the same time, gamification has been increasingly applied to enhance engagement in language learning. By integrating elements such as points, rewards, and feedback, gamification can increase learner participation and motivation (X. Li et al., 2022; Zhang & Hasim, 2023). In writing instruction, however, gamification is generally considered an indirect support mechanism, influencing engagement and readiness rather than directly improving writing skills (Wang, 2023).

One widely used gamified platform is ClassDojo, which provides real-time feedback and reward systems to reinforce learner participation (Corser et al., 2025). While ClassDojo has been commonly used for classroom management and engagement, it has rarely been examined as a pedagogical tool integrated into specific instructional stages. In particular, its role within warm-up activities as a pre-writing support mechanism remains underexplored.

This gap is particularly relevant in Vietnamese EFL contexts, where learners often face difficulties in writing due to both linguistic limitations and low motivation (Nguyen, 2021; Whitty et al., 2022). In private language centers, where instructional time is limited and maintaining engagement is essential, the need for effective pre-writing support becomes more pronounced. Although digital tools are widely adopted, their use has largely focused on classroom management rather than writing preparation (Phan, 2024).

Addressing this gap, the present study investigates the integration of ClassDojo into warm-up activities in EFL writing lessons at a Vietnamese language center. Rather than positioning ClassDojo as a direct instructional tool for teaching writing, the study conceptualizes it as an indirect pedagogical support embedded within the pre-writing stage. Through features such as points, feedback, and visual rewards, ClassDojo is expected to enhance learners' engagement, readiness, and motivation before writing tasks, which may in turn be associated with improvements in writing performance.

Importantly, this study does not assume a direct causal relationship between the use of ClassDojo and writing outcomes. Instead, writing performance is interpreted as a context-dependent result shaped by multiple factors, including learner engagement and readiness. In addition to examining immediate performance outcomes, the study also considers whether such improvements are sustained over time, thereby providing a more comprehensive understanding of instructional impact.

The study is guided by the following research questions:

1. To what extent does the integration of ClassDojo into warm-up activities affect EFL learners' writing performance?

2. What are learners' perceptions of using ClassDojo in warm-up activities to support their writing performance?

By focusing on both performance outcomes and learner perceptions, this study contributes to the literature by situating gamification within the pre-writing stage of instruction. The findings are expected to provide practical implications for integrating gamified tools into warm-up activities, particularly in EFL contexts where enhancing learner readiness and engagement is a key instructional concern.

2. Literature Review

2.1 Writing Performance in EFL Contexts

Writing performance in EFL contexts is commonly defined as learners' observable written output produced under specific task conditions and evaluated using explicit criteria (Weigle, 2002; Shaw & Weir, 2007). Unlike writing ability, which reflects underlying competence, writing performance is context-bound and measurable, making it suitable for empirical investigation.

In standardized frameworks such as those of Cambridge Assessment English, writing performance is typically assessed across dimensions including content, organization, and language use (Cambridge University Press & Assessment, 2024). This multidimensional approach enhances reliability and allows for criterion-referenced evaluation. However, performance is not solely determined by linguistic competence. It reflects the interaction between learners' knowledge, task demands, and contextual conditions (Hyland, 2003; Weigle, 2002).

Importantly, recent studies emphasize that writing performance is also influenced by pre-task conditions, including learners' engagement and readiness (Fleckenstein et al., 2024; Zenouzagh et al., 2025). Therefore, improvements in writing outcomes should be interpreted as context-dependent rather than direct results of a single instructional intervention. In this study, writing performance is examined both in terms of immediate outcomes and potential sustained effects over time.

2.2 Cognitive and Affective Demands in EFL Writing

Writing is widely recognized as a cognitively demanding activity involving multiple recursive processes such as planning, translating, and reviewing (Flower & Hayes, 1981). Among these, planning is particularly critical in the early stage, as learners must generate ideas and organize content before producing written text.

From the perspective of Cognitive Load Theory (Sweller, 1988), writing tasks impose a high cognitive load because learners must simultaneously manage idea development, linguistic encoding, and task requirements. This challenge is especially pronounced among lower-proficiency learners, who may struggle to initiate writing due to limited lexical resources (Alzubi & Nazim, 2024).

In addition to cognitive constraints, affective factors such as anxiety and low confidence play a significant role. Writing anxiety has been consistently identified as a barrier to effective performance in EFL contexts (Akhtar et al., 2019; Rasool et al., 2023; Sun et al., 2024). Learners experiencing anxiety often demonstrate reduced willingness to participate and lower task engagement.

Importantly, both cognitive and affective challenges tend to emerge before writing begins, affecting learners' readiness to engage in the task. This highlights the importance of instructional support at the pre-writing stage, where learners can be prepared cognitively and emotionally for writing tasks.

2.3 Warm-up Activities and Writing Readiness

Warm-up or pre-writing activities are traditionally used as introductory stages in language lessons. However, recent research suggests that they play a more substantial role as a mechanism for enhancing writing readiness (Yuk & Yunus, 2021; Mahdi et al., 2025).

Warm-up activities typically involve low-stakes, interactive tasks such as brainstorming, vocabulary activation, or short discussions. From a cognitive perspective, these

activities facilitate idea generation and lexical retrieval, enabling learners to approach writing tasks with greater clarity (Mahdi et al., 2025). From an affective perspective, they reduce anxiety and increase confidence by providing a gradual transition into writing (Rasool et al., 2023). Behaviorally, they promote participation and establish early engagement.

These functions align with the concept of writing readiness, which includes cognitive readiness (availability of ideas), affective readiness (confidence), and behavioral readiness (willingness to engage). Unlike engagement, which occurs during task performance, readiness operates at the pre-task stage and influences how learners initiate writing.

Despite their importance, warm-up activities are often minimized in time-constrained classrooms (Nguyen-Viet & Nguyen-Viet, 2023). As a result, learners may begin writing tasks without sufficient preparation, which can negatively affect performance.

In this study, warm-up activities are positioned as the primary instructional stage where readiness can be enhanced. However, consistent with the research design, they are not assumed to directly improve writing performance but to function as an enabling condition that may indirectly support performance outcomes.

2.4 Gamification as an Indirect Support Mechanism

Gamification refers to the integration of game elements such as points, rewards, and feedback into non-game learning environments (X. Li et al., 2022). In EFL education, gamification has been widely associated with increased engagement, motivation, and participation (Zhang & Hasim, 2023; Chan & Lo, 2024).

Empirical studies indicate that gamified learning environments can enhance behavioral engagement, emotional involvement, and learning persistence (Abdulbaki et al., 2025; Boudadi et al., 2024). However, its direct impact on academic performance, particularly writing, remains less conclusive.

In writing instruction, gamification is more appropriately understood as an indirect support mechanism. Rather than directly improving writing skills, it influences mediating factors such as engagement, motivation, and anxiety reduction (Wang, 2023; Núñez-Pacheco et al., 2023). These factors may, in turn, create more favorable conditions for writing.

This distinction is essential for maintaining academic rigor and avoiding overgeneralized claims about effectiveness.

2.5 ClassDojo in Warm-up Activities

ClassDojo is a widely used digital platform that incorporates gamified features such as real-time feedback, point systems, and visual rewards. It is primarily designed for classroom management, allowing teachers to reinforce participation and monitor behavior (Corser et al., 2025; Pham, 2021).

Previous research has shown that ClassDojo and other gamified digital platforms can enhance learner engagement, participation, and classroom interaction (Sotos-Martínez et al., 2023; Phan, 2024). However, research on ClassDojo has largely focused on behavior management rather than instructional processes.

Importantly, there is limited research examining how ClassDojo can be integrated into specific instructional stages, particularly warm-up activities in writing lessons. This represents a key gap in the literature.

In the present study, ClassDojo is positioned explicitly within the warm-up stage. Its features are used to structure short, interactive activities that provide immediate feedback, encourage participation, and create a motivating environment before writing tasks. Through this integration, ClassDojo is expected to enhance learners' engagement, readiness, and motivation at the pre-writing stage.

Consistent with the study's conceptual framework, ClassDojo is not treated as a direct writing instruction tool. Instead, it functions as an indirect pedagogical support, facilitating the conditions under which writing tasks are performed.

2.6 Learner Perceptions: Technology Acceptance Model (TAM) and Self-Determination Theory (SDT) Perspectives

Learner perceptions are critical in understanding how digital tools function in educational contexts. In this study, perceptions are examined through the combined perspectives of the TAM and SDT.

TAM (Davis, 1989) explains technology acceptance through two key constructs: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). These constructs influence learners' willingness to adopt and engage with digital tools (Davis, 1989).

Complementing TAM SDT (Ryan & Deci, 2000) explains motivation in terms of psychological needs such as competence and autonomy. Gamified features such as feedback and rewards may support these needs, thereby enhancing intrinsic motivation.

Building on these frameworks, the present study examines learner perceptions across four constructs: perceived ease of use, perceived usefulness, engagement and readiness, and motivation. These constructs are operationalized through the student questionnaire and directly address the second research question.

2.7 Research Gap and Conceptual Positioning

Despite the growing body of research on gamification and digital tools in EFL education, several gaps remain.

First, existing studies have primarily focused on engagement and motivation, with limited attention to writing performance as a measurable outcome (Zhang & Hasim, 2023; Wang, 2023). Second, ClassDojo has rarely been examined within specific instructional stages, particularly the warm-up phase. Third, there is limited research integrating performance outcomes with learner perceptions within a unified framework.

These gaps are particularly evident in Vietnamese EFL contexts, where writing instruction remains challenging and the pedagogical use of gamified tools in pre-writing stages has not been systematically explored (Nguyen, 2021; Whitty et al., 2022).

To address these gaps, the present study proposes an indirect mechanism in which ClassDojo is integrated into warm-up activities to enhance learners' engagement, readiness, and motivation prior to writing tasks. These factors may, in turn, be associated with improvements in writing performance, including potential sustained effects over time. At the same time, learners' perceptions of this integration are examined through perceived usefulness, perceived ease of use, engagement and readiness, and motivation.

This conceptual positioning provides the theoretical foundation for the present study, which examines the integration of ClassDojo into warm-up activities as an indirect support mechanism for EFL learners' writing performance within a specific instructional context.

3. Methodology

The present study employed a quasi-experimental mixed-methods design to examine the integration of ClassDojo into warm-up activities and its association with EFL learners' writing performance. This design combined quantitative and qualitative approaches to capture both measurable changes in writing outcomes and learners' perceptions of the intervention.

3.1 Research Design and Participants

The quantitative component followed a non-equivalent control group design, as intact classes were used and random assignment was not feasible in the instructional context. Writing performance was measured at three points: pre-test, immediate post-test, and delayed post-test (two weeks later), allowing the study to examine both improvement and short-term retention.

The study was conducted at Nam Do Language Center (Ho Chi Minh City, Vietnam). Four KET-level classes participated. The initial sample included 86 learners; after removing incomplete cases, the final sample consisted of 73 learners (38 experimental, 35 control). Both groups followed the same curriculum, with the only difference being the integration of ClassDojo into the experimental group's warm-up activities.

3.2 Instruments

Writing performance was assessed using three tests aligned with the Cambridge A2 Key (KET) format. Each test included a communicative task and a guided writing task. Scripts were rated independently by two raters using the Cambridge analytic rubric (Content, Organisation, and Language), and the average score was used.

Inter-rater consistency was examined using Pearson's correlation coefficient. The correlations between the two raters' scores were consistently strong across the pre-test, immediate post-test, and delayed post-test, ranging from $r = .861$ to $r = .916$, all $p < .001$. Therefore, the average score of the two raters was used for further statistical analysis. The detailed inter-rater correlation coefficients are presented in Table 1.

Table 1
Inter-Rater Correlations of Writing Test Scores

Test	Group	r	p
Pre-test	Experimental	.899	< .001
Pre-test	Control	.916	< .001
Post-test	Experimental	.861	< .001
Post-test	Control	.889	< .001
Delayed post-test	Experimental	.871	< .001
Delayed post-test	Control	.908	< .001

Learners' perceptions were collected through a bilingual questionnaire administered to the experimental group. The instrument measured four constructs: perceived ease of use,

perceived usefulness, engagement and readiness, and motivation. In addition, semi-structured interviews were conducted with ten learners to gain deeper insights into their experiences.

3.3 Intervention Procedure

The intervention lasted twelve weeks, with three sessions per week. In the experimental group, ClassDojo was integrated into the warm-up stage (10–15 minutes) of each writing lesson. Activities included brainstorming, vocabulary recall, and short interactive tasks, supported by real-time feedback and participation points.

The purpose of this integration was to enhance learners' engagement and readiness prior to writing, rather than to directly teach writing. The control group followed the same instructional content without ClassDojo.

3.4 Data Collection and Analysis

Data collection followed a fixed and systematic sequence. Initially, both groups completed the pre-test to establish baseline writing performance. This was followed by the intervention period, after which the immediate post-test was administered. Subsequently, the questionnaire and semi-structured interviews were conducted with the experimental group to explore learners' perceptions. Finally, a delayed post-test was administered to both groups to examine the stability of performance over time. This structured procedure allowed the study to capture both changes in writing performance and learners' perceptions of the intervention.

Quantitative data were analyzed using descriptive statistics, paired-samples and independent-samples t-tests, and effect sizes (Cohen's d). Questionnaire data were analyzed descriptively, with internal consistency assessed using Cronbach's alpha.

Qualitative data from interviews were analyzed using thematic analysis to identify patterns related to learners' perceptions of engagement, readiness, and usefulness. Findings from both strands were integrated at the interpretation stage.

3.5 Ethical Considerations

Participants were informed about the study and participated voluntarily. All data were anonymized, and participation had no impact on academic evaluation.

4. Results

This section presents the findings in relation to the two research questions. It first reports writing performance across the experimental and control groups and then examines learners' perceptions of the ClassDojo-supported warm-up activities. These findings provide the basis for the subsequent discussion of writing performance, engagement, readiness, and motivation.

4.1. Writing Performance Outcomes

4.1.1. Baseline comparison

Before examining post-intervention outcomes, the pre-test writing scores of the experimental and control groups were compared. Table 2 presents the baseline comparison.

Table 2
Baseline Comparison of Pre-test Writing Scores

Group	N	Mean	SD	t	p

Experimental	38	15.53	5.04	0.725	.471
Control	35	14.70	4.66		

Table 2 indicates that the mean score of the experimental group was slightly higher than that of the control group. However, the difference was not statistically significant (p = .471), indicating that no statistically significant difference in pre-test writing performance was detected between the two groups. This provides a reasonable baseline for subsequent comparisons.

4.1.2. Within-group development

To investigate the developmental changes over time, writing scores of both groups were compared at pre-test, post-test, and delayed post-test phases. Table 3 provides the results.

Table 3

Within-Group Development of Writing Scores

Group	**Pre-test**	**Post-test**	**Delayed Post-test**	**Gain**
Experimental	15.53	22.36	22.58	6.83
Control	14.70	18.78	18.00	4.07

As Table 3 shows, both groups improved from pre-test to post-test, although the magnitude and maintenance of the gains differed. The experimental group showed substantial improvement from the pre-test to the post-test (gain = 6.83), and this performance was maintained at the delayed post-test. The control group also improved from pre-test to post-test (gain = 4.07), but its mean score declined slightly at the delayed stage, indicating less stable performance over time.

4.1.3. Between-group comparison

To examine between-group differences further, writing performance was compared at the post-test and delayed post-test. Table 4 presents the results.

Table 4

Between-Group Comparison of Writing Performance

Stage	**Group**	**Mean**	**SD**	**t**	**p**
Post-test	Experimental Group	22.36	4.11	3.426	.001
	Control Group	18.78	4.82		
Delayed	Experimental Group	22.58	3.44	4.469	< .001
	Control Group	18.00	5.20		

The results indicate significant differences between the two groups at both the post-test ($p = .001$) and delayed post-test ($p < .001$). The experimental group consistently outperformed the control group, demonstrating greater immediate improvement and more stable short-term retention. The wider between-group difference at the delayed stage further suggests that the observed performance advantage was maintained over the short term, whereas the control group showed a slight decline.

4.1.4. Effect size

In order to gain a better insight into the practical importance of these findings, effect sizes were computed and are shown in Table 5.

Table 5

Effect Sizes for Writing Performance

Comparison	**Cohen's d**
Experimental (Pre–Post)	1.93
Control (Pre–Post)	1.63
Between groups (Post-test)	0.80

Both groups showed substantial improvement; however, the effect was stronger in the experimental group. The between-group effect size indicates that the post-test difference was not only statistically significant but also educationally meaningful.

4.2. Learners' Perceptions

In addition to writing performance, the study also examined learners' perceptions of the ClassDojo-supported warm-up activities. The reliability of the questionnaire constructs is provided in Table 6.

Table 6

Reliability of Questionnaire Constructs

Construct	**Items**	**Cronbach's Alpha**
Ease of Use	6	.811
Usefulness	6	.852
Engagement & Readiness	6	.847
Motivation	6	.887

Internal consistency of all constructs was acceptable, indicating that the questionnaire was a reliable instrument for measuring learners' perceptions.

Learners generally reported positive perceptions of ClassDojo, particularly in terms of ease of use and usefulness. The platform was perceived as intuitive, allowing students to participate in activities with minimal difficulty. In terms of usefulness, learners indicated that the warm-up activities supported idea generation, vocabulary recall, and better understanding of writing tasks before beginning.

In addition, students reported increased engagement and readiness during the warm-up phase. The use of real-time feedback and point-based rewards encouraged active participation and helped create a more supportive learning environment. Learners also reported feeling more confident and less anxious when approaching writing tasks, and they perceived the gamified features as motivating, contributing to a more positive overall learning experience.

Translated interview data further supported the questionnaire findings. One learner stated, “When I used points, I really paid more attention.” Another student explained, “Normally I am afraid because I do not know what to write, but with the warm-up questions and points, I was more focused and felt a bit more confident.” A third participant added, “I began the writing task faster than I used to.” These responses suggest that ClassDojo-supported warm-up activities helped learners become more attentive, confident, and ready before writing.

5. Discussion

This section interprets the findings in relation to the research questions and situates them within existing theoretical and empirical literature. The discussion focuses on the role of ClassDojo-supported warm-up activities as an indirect mechanism influencing writing performance and learners’ perceptions.

5.1 ClassDojo as an Indirect Support for Writing Performance

In this study, ClassDojo did not function as a direct instructional tool for writing. It was not used to teach writing structure, provide corrective feedback, or improve linguistic accuracy directly; instead, it was integrated into the warm-up stage prior to writing tasks. Within this instructional design, the experimental group demonstrated greater improvement and more stable short-term retention than the control group.

One explanation for this finding is that ClassDojo-supported warm-up activities enhanced pre-writing conditions, particularly in terms of idea activation, readiness, and engagement. From a cognitive perspective, these activities may have reduced the processing demands during writing. According to the cognitive process model of writing (Flower & Hayes, 1981), planning is a critical stage in which learners generate and organize ideas before text production. By activating relevant vocabulary and ideas during the warm-up phase, learners were better prepared to begin writing, thereby reducing the need to manage multiple processes simultaneously.

This interpretation is further supported by Cognitive Load Theory (Sweller, 1988), which suggests that reducing cognitive load can facilitate task performance. When learners enter a writing task with pre-activated ideas and vocabulary, they may experience less cognitive overload, allowing them to allocate more resources to organizing and expressing their ideas.

In addition, the gamified features of ClassDojo may have contributed to increased engagement and attention during the pre-writing stage. Previous research has shown that gamification can enhance learner participation and emotional involvement (C. Li et al., 2024; Zhang & Hasim, 2023). In the present study, this increased engagement may have created more favorable conditions for writing, which may explain the observed differences in performance between groups.

Taken together, these findings suggest that the contribution of ClassDojo lies not in directly improving writing skills, but in enhancing the conditions under which writing occurs. This supports the view that writing performance is influenced by multiple interacting factors, including pre-task readiness and engagement.

5.2 The Role of Warm-up Activities in Writing Instruction

The findings also highlight the pedagogical importance of the warm-up stage in writing lessons. Rather than functioning as a routine or transitional activity, warm-up can be conceptualized as a structured pre-writing phase that supports learners' cognitive and emotional preparation.

In this study, warm-up activities facilitated idea generation, vocabulary recall, and increased focus, which are essential for effective writing initiation. This aligns with previous research suggesting that pre-writing support can improve learners' readiness and reduce hesitation at the early stage of writing (Yuk & Yunus, 2021; Mahdi et al., 2025).

Importantly, the results suggest that writing development begins before the act of drafting. Learners need to activate relevant knowledge, organize initial ideas, and overcome psychological barriers before they can engage effectively in writing tasks. Therefore, instructional approaches that support this preparatory phase may have a meaningful impact on overall writing performance.

This perspective extends traditional models of writing instruction by emphasizing the importance of pre-task conditions, particularly in EFL contexts where learners often struggle to initiate writing.

5.3 Contextual Suitability of ClassDojo

The observed benefits of ClassDojo in this study appear to be closely related to the characteristics of the instructional context. First, the platform provides immediate feedback and visible participation cues, which help maintain learners' attention at the beginning of the lesson. This is particularly important in warm-up activities, where initial engagement sets the tone for the rest of the lesson.

Second, the use of gamified elements such as points and visual rewards creates a more interactive and motivating classroom environment compared to traditional warm-up practices. This supports previous findings that gamification can enhance classroom interaction and learner involvement (Chan & Lo, 2024).

Furthermore, ClassDojo appeared particularly suitable for the KET-level learners in this study because of its simplicity and accessibility. The platform did not require advanced technical skills, allowing learners to participate with minimal technological demands. This alignment between tool design and learner characteristics may have contributed to the observed improvements in engagement and readiness.

These findings suggest that the success of educational technology depends not only on its features but also on its fit with the instructional context and learner needs.

5.4 Learner Perceptions: TAM and SDT Perspectives

The positive perceptions reported by learners can be interpreted through TAM (Davis, 1989). Learners perceived ClassDojo as both useful and easy to use, which likely contributed to their willingness to engage with the platform. Perceived usefulness was reflected in learners'

reports that warm-up activities supported idea generation and task understanding, while perceived ease of use was associated with the platform's intuitive design.

In addition, the findings can be explained through SDT (Ryan & Deci, 2000). The use of real-time feedback and rewards may have supported learners' sense of competence and engagement, potentially supporting learners' motivation. This is consistent with previous research showing that gamified environments can positively influence learners' motivation and emotional involvement (Abdulbaki et al., 2025).

Importantly, the study highlights engagement and readiness as potential explanatory factors rather than statistically tested mediators. Learners reported feeling more prepared, confident, and less anxious before writing, suggesting that the intervention was associated with changes in their pre-task state. This supports the view that learner perceptions, engagement, and performance outcomes may be closely related.

5.5 Pedagogical and Theoretical Implications

Pedagogically, the findings suggest that teachers should treat warm-up activities as structured pre-writing support rather than routine introductory tasks. In EFL writing lessons, warm-up activities can be designed to activate learners' ideas, support vocabulary recall, and increase readiness before the main writing task. Simple gamified tools such as ClassDojo may help teachers create a more interactive and motivating pre-writing environment without requiring complex technological integration. By using features such as immediate feedback and participation rewards, teachers can encourage learners to become more focused, confident, and prepared before writing.

Theoretically, this study contributes to the understanding of writing instruction by emphasizing the role of pre-task conditions in shaping writing performance. The findings support the view that engagement and readiness are important factors in the writing process, particularly for lower-proficiency EFL learners. In addition, the study contributes to the literature on gamification by showing that gamified tools may operate as indirect support mechanisms. Rather than directly improving writing skills, ClassDojo may support learners' motivation, emotional readiness, and participation, thereby creating more favorable conditions for writing performance.

6. Conclusion

This study examined the integration of ClassDojo into warm-up activities as a form of pedagogical support for EFL writing instruction in a Vietnamese language center. Rather than functioning as a direct instructional tool, ClassDojo was embedded in the pre-writing phase to enhance learners' engagement, readiness, and motivation prior to writing tasks.

The findings indicated that the experimental group demonstrated greater improvement in writing performance than the control group. This improvement was observed at the immediate post-test and was maintained at the delayed post-test, suggesting more stable short-term retention. In addition, learners reported generally positive perceptions of the ClassDojo-supported warm-up activities, particularly in terms of perceived ease of use, perceived usefulness, engagement and readiness, and motivation.

Taken together, these results suggest that ClassDojo can support writing development indirectly by enhancing the conditions under which writing occurs. The platform may contribute to a more supportive pre-writing environment by facilitating idea generation, increasing learner engagement, and improving readiness to write.

From a pedagogical perspective, the findings highlight the importance of the warm-up phase in writing instruction. Teachers may benefit from designing structured pre-writing activities that support both cognitive and affective preparation. In this context, simple gamified tools such as ClassDojo offer a practical and accessible way to enhance learner participation without introducing unnecessary complexity.

Several limitations should be acknowledged. The study employed a quasi-experimental design using intact classes, which limits the strength of causal inferences. In addition, the research was conducted within a single language center with KET-level learners, which may affect the generalizability of the findings. The relatively short duration of the intervention should also be considered when interpreting the results.

Future research may explore the effectiveness of similar interventions across different proficiency levels and over longer periods of time. Further studies could also investigate other forms of technology-supported warm-up activities and their potential impact on writing performance in EFL contexts.

Overall, the study suggests that even modest adjustments in instructional design, particularly at the pre-writing stage, can contribute to meaningful differences in learner outcomes. Practical and user-friendly tools such as ClassDojo may offer valuable support for enhancing engagement and writing development in EFL classrooms.

Acknowledgments

The authors would like to express sincere gratitude to the management and teaching staff at Nam Do Language Center for their support in facilitating this study. Special thanks are extended to the students who participated in the research for their time and cooperation. The authors also appreciate the valuable guidance and feedback from academic supervisors and colleagues throughout the research and writing process.

References

Abdulbaki, S., Khasawneh, M. A. S., & Tashtoush, M. A. (2025). The effectiveness of gamified learning environments in promoting grammar mastery in Jordanian secondary school EFL learners. *International Journal of Innovative Research and Scientific Studies*, *8*(2), 3375–3386. https://doi.org/10.53894/ijirss.v8i2.6013

Akhtar, R., Hassan, H., Saidalvi, A., & Hussain, S. (2019). A systematic review of the challenges and solutions of ESL students' academic writing. *International Journal of Engineering and Advanced Technology*, *8*(5c), 1169–1171. https://doi.org/10.35940/ijeat.E1164.0585C19

Alzubi, A. A. F., & Nazim, M. (2024). Students' intrinsic motivation in EFL academic writing: Topic-based interest in focus. *Heliyon*, *10*(1), e24169. https://doi.org/10.1016/j.heliyon.2024.e24169

Boudadi, N. A., Gutiérrez-Colón, M., & Usart, M. (2024). A gamified learning environment (Moodle) to enhance English language learning at university level. *International Journal of Instruction*, *17*(4), 483–502. https://doi.org/10.29333/iji.2024.17427a

Cambridge University Press & Assessment. (2024). *Assessing writing for Cambridge English Qualifications: A guide for teachers – A2 Key for Schools*.

Chan, S., & Lo, N. (2024). Enhancing EFL/ESL instruction through gamification: A comprehensive review of empirical evidence. *Frontiers in Education*, *9*, 1–19. https://doi.org/10.3389/feduc.2024.1395155

Corser, K., Manolev, J., & Danby, S. (2025). Problematising ClassDojo as a digital tool for behaviour management and home-school communication. *Learning, Media and Technology*, 1–17. https://doi.org/10.1080/17439884.2025.2553184

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Fleckenstein, J., Jansen, T., Meyer, J., Trüb, R., Raubach, E. E., & Keller, S. D. (2024). How am I going? Behavioral engagement mediates the effect of individual feedback on writing performance. *Learning and Instruction*, *93*. https://doi.org/10.1016/j.learninstruc.2024.101977

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.2307/356600

Hyland, K. (2003). *Second language writing*. Cambridge University Press. https://www.cambridge.org/core/books/second-language-writing/B55667B789336DF8690E651E5111E6A4

Li, C., Wei, L., & Lu, X. (2024). Task complexity and L2 writing performance of young learners: Contributions of cognitive and affective factors. *Modern Language Journal*, *108*(3), 741–770. https://doi.org/10.1111/modl.12954

Li, X., Xia, Q., Chu, S. K. W., & Yang, Y. (2022). Using gamification to facilitate students' self-regulation in e-learning: A case study on students' L2 English learning. *Sustainability*, *14*(12), 7008. https://doi.org/10.3390/su14127008

Mahdi, H. S., Milad, M., & Saeed, A. T. (2025). Evaluating the effectiveness of AI tools across the essay writing process: A comparative study of pre-writing, drafting, and post-

writing stages. *International Journal of Computer-Assisted Language Learning and Teaching*, *15*(1), 1–20. https://doi.org/10.4018/IJCALLT.391354

Nguyen, T. H. (2021). Improving writing skills with systemic functional linguistic approach: The case of Vietnamese EFL students. *World Journal of English Language*, *11*(2), 71–83. https://doi.org/10.5430/wjel.v11n2p71

Nguyen-Viet, B., & Nguyen-Viet, B. (2023). Enhancing satisfaction among Vietnamese students through gamification: The mediating role of engagement and learning effectiveness. *Cogent Education*, *10*(2), 1–19. https://doi.org/10.1080/2331186X.2023.2265276

Núñez-Pacheco, R., Vidal, E., Castro-Gutierrez, E., Turpo-Gebera, O., Barreda-Parra, A., & Aguaded, I. (2023). Use of a gamified platform to improve scientific writing in engineering students. *Education Sciences*, *13*(12). https://doi.org/10.3390/educsci13121164

Pham, T. H. M. (2021). *Using ClassDojo to manage young learners' classroom behaviors and improve parent-teacher communication in the Vietnamese context*. Ho Chi Minh City Open University.

Phan, N. T. T. (2024). Attitudes of English majors towards the use of the Quizizz platform in English language learning. *Ho Chi Minh City Open University Journal of Science - Social Sciences*, *14*(4), 69–83. https://doi.org/10.46223/HCMCOUJS.soci.en.14.4.2951.2024

Rasool, U., Qian, J., & Aslam, M. Z. (2023). An investigation of foreign language writing anxiety and its reasons among pre-service EFL teachers in Pakistan. *Frontiers in Psychology*, *13*.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Shaw, S. D., & Weir, C. J. (2007). *Examining writing: Research and practice in assessing second language writing*. Cambridge University Press.

Sotos-Martínez, V. J., Tortosa-Martínez, J., Baena-Morales, S., & Ferriz-Valero, A. (2023). Boosting student's motivation through gamification in physical education. *Behavioral Sciences*, *13*(2), 165. https://doi.org/10.3390/bs13020165

Sun, J., Motevalli, S., & Chan, N. N. (2024). Exploring writing anxiety during writing process: An analysis of perceptions in Chinese English as a foreign language (EFL) learners. *Qualitative Research in Education*, *13*(2), 149–164. https://doi.org/10.17583/qre.12938

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Wang, Y.-C. (2023). Gamification in a news English course. *Education Sciences*, *13*(1), 90. https://doi.org/10.3390/educsci13010090

Weigle, S. C. (2002). *Assessing writing* (J. C. Alderson & L. F. Bachman, Eds.). Cambridge University Press.

Whitty, L., Parkinson, J., & Pham, H. T. P. (2022). Can and could in academic writing: A corpus-driven comparison of English L1 and Vietnamese EFL students. *The Journal of Asia TEFL*, *19*(1), 93–108. https://doi.org/10.18823/asiatefl.2022.19.1.6.93

Yuk, A. C. K., & Yunus, M. M. (2021). Using peer-modo feedback at the pre-writing stage to improve year 4 pupils' writing performance. *Journal of Education and E-Learning Research*, *8*(1), 116–124. https://doi.org/10.20448/journal.509.2021.81.116.124

Zenouzagh, Z. M., Admiraal, W., & Saab, N. (2025). Empowering student engagement: The dynamics of learner traits in digital feedback environments. *Journal of Computing in Higher Education*. https://doi.org/10.1007/s12528-025-09459-z

Zhang, S., & Hasim, Z. (2023). Gamification in EFL/ESL instruction: A systematic review of empirical research. *Frontiers in Psychology*, *13*, 1030790. https://doi.org/10.3389/fpsyg.2022.1030790

#79 EXPLORING THE EFFECTS OF AI-MEDIATED LOW-STAKES PRACTICE ON SPEAKING ANXIETY IN VIETNAMESE EFL CLASSROOMS: A QUASI-EXPERIMENTAL STUDY

Minh Hang Le[1]

Abstract

Speaking anxiety continues to constrain learners' participation and oral performance in English-as-a-Foreign-Language (EFL) classrooms. Although recent work has examined the use of artificial intelligence (AI) in language learning, its role in addressing affective factors remains insufficiently investigated. This study examines the effects of AI-mediated low-stakes practice on speaking anxiety among Vietnamese university students. A quasi-experimental design was implemented with two intact classes of 64 second-year English-major students at a Vietnamese university. The experimental group engaged in AI-supported speaking activities, including guided role-plays and repeated practice using conversational AI tools over a six-week period, while the control group followed conventional speaking instruction. Data were collected through pre- and post-tests using a speaking anxiety questionnaire adapted from the Foreign Language Classroom Anxiety Scale (FLCAS) and an analytic speaking rubric. The results indicate that students in the experimental group reported significantly lower levels of speaking anxiety and higher levels of confidence compared to those in the control group. Student feedback further suggests that AI-mediated practice enabled repeated, low-pressure exposure and reduced concerns about negative evaluation. The findings provide empirical support for the integration of AI-mediated low-stakes activities in speaking instruction as a means of addressing affective constraints and supporting learner participation in EFL contexts.

Keywords: AI-mediated learning, speaking anxiety, low-stakes practice, EFL classrooms, quasi-experimental design

1. Introduction

Speaking is widely regarded as the most anxiety-provoking skill in foreign language learning (Horwitz et al., 1986; Young, 1991). Horwitz et al. (1986) identified three interrelated components of foreign language anxiety (FLA): communication apprehension, fear of negative evaluation, and test anxiety. Subsequent factor-analytic work confirmed the distinctiveness of speaking anxiety within the broader FLA construct (Aida, 1994). Of these three components, fear of negative evaluation has been most consistently associated with speaking reticence in EFL contexts (MacIntyre & Gardner, 1994; Teimouri et al., 2019; Young, 1991). In English-as-a-Foreign-Language (EFL) settings, speaking anxiety frequently undermines participation and overall oral performance (Liu & Jackson, 2008; Woodrow, 2006). Meta-analytic evidence confirms that the negative relationship between anxiety and L2 achievement is robust, with speaking consistently the most susceptible domain (Teimouri et al., 2019).

[1] *Lecturer, Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

Vietnamese EFL classrooms present a particularly acute instance. Large class sizes, examination-oriented curricula, and Confucian-heritage norms that privilege face preservation interact synergistically to compound speaking anxiety (Bao, 2013). Truong and Ho (2022), in a study of 297 Vietnamese freshmen, confirmed that speaking anxiety remains pervasive, with communication apprehension, fear of negative evaluation, and test anxiety all contributing to students' reluctance to speak. Tran and Moni (2015) further showed that Vietnamese learners' reticence was mediated by collectivist norms of passive reception. Despite the well-documented severity of this problem, no systematic attempt has been made to investigate whether emerging AI technologies can alleviate speaking anxiety in Vietnamese higher education. This absence is particularly problematic given that Vietnam's Ministry of Education and Training has actively promoted technology integration in foreign language teaching, yet empirical evidence to guide AI-mediated pedagogical practices in this context remains unavailable.

The motivation for this study arose from a recurring observation in Vietnamese university speaking classrooms: students who demonstrate strong grammatical knowledge and reading comprehension in written assessments become visibly anxious and largely silent when asked to speak. In typical large speaking classes, it is not uncommon for only a handful of students to voluntarily contribute to oral activities, while the majority resort to rehearsed scripts or avoid eye contact to escape being called upon. This persistent gap between what students know and what they are willing to say motivated the present investigation into whether AI-mediated practice could offer a way to break the cycle of silence.

Against this backdrop, voice-enabled AI platforms now allow learners to practice speaking without the social pressure of face-to-face interaction (Fathi et al., 2024; Huang et al., 2022). For the purposes of this study, low-stakes practice is operationally defined as speaking activities that: (a) are not formally graded, (b) do not involve real-time performance before a peer or instructor audience, and (c) permit self-paced repetition without social consequence. This definition draws on the desirable difficulties framework (Bjork & Bjork, 2011) and recent AI-mediated speaking research (Fryer et al., 2019; Fathi et al., 2024).

The theoretical rationale draws on Krashen's (1982) affective filter hypothesis and Horwitz et al.'s (1986) tripartite model. Despite acknowledged limitations regarding empirical testability (McLaughlin, 1987), the affective filter remains a productive heuristic, and its core prediction—that reducing anxiety facilitates acquisition—continues to receive empirical support (Teimouri et al., 2019). AI-mediated practice may lower this filter by targeting two of the three FLA components. Fear of negative evaluation is reduced because the AI interlocutor is non-judgmental and does not relay performance information to peers. Communication apprehension is attenuated because private, repeatable practice removes real-time audience pressure. Empirical support for these mechanisms comes from Zhang et al. (2024), who found reduced anxiety and increased enjoyment among Chinese EFL learners using AI chatbots, and Fathi et al. (2024), who reported improved speaking skills in an AI-mediated experimental condition. Test anxiety, tied to formal assessment contexts, is less directly addressed.

A necessary caveat, however, is that AI-mediated interaction is not anxiety-neutral. Speech recognition failures, unnatural turn-taking, and the sheer unfamiliarity of talking to a machine could partially offset the benefits of a non-evaluative practice space (El Shazly, 2021). Recent findings suggest such frustrations diminish as familiarity develops (Chen et al., 2025; Fathi et al., 2024).

The empirical base, though growing rapidly, remains thin and geographically uneven. Fathi et al. (2024), in a true experimental study with 40 Iranian EFL learners, found significant

improvements in speaking skills, though the small sample limits statistical power. Zhang et al. (2024) reported reduced anxiety among 99 Chinese students with moderate effect sizes. Chen et al. (2025) found significant anxiety reductions with 71 Chinese learners over eight weeks. Wu et al. (2026) demonstrated reduced anxiety among 75 Indonesian students. These studies span East Asian, Middle Eastern, and Southeast Asian contexts, but none has examined Vietnamese higher education. A systematic search of Scopus, Web of Science, and ERIC in January 2025 using ("artificial intelligence" OR "chatbot" OR "AI") AND ("speaking anxiety" OR "oral anxiety" OR "FLSA") AND ("Vietnam*") in titles, abstracts, and keywords yielded zero results.

The present study was designed to fill this void. The inclusion of oral performance as a dependent variable is theoretically motivated: if AI-mediated practice lowers the affective filter, learners should show greater willingness to attempt output and increased engagement with input, facilitating gains in fluency and interactional competence (Krashen, 1982). Furthermore, repeated practice may develop procedural knowledge through automatization (DeKeyser, 2007). By examining anxiety sub-components alongside oral performance and learner perceptions, this study contributes not only contextual evidence from Vietnam but also a more fine-grained understanding of how AI-mediated low-stakes practice may influence affective and communicative dimensions of EFL speaking development. Accordingly, this study pursues three specific objectives: (1) to measure the effect of AI-mediated low-stakes practice on total and sub-component speaking anxiety; (2) to assess whether anxiety reduction translates into measurable improvements in oral performance; and (3) to explore students' lived experiences with AI-mediated practice through qualitative inquiry. The study is guided by the following research questions:

RQ1: To what extent does AI-mediated low-stakes speaking practice reduce foreign language speaking anxiety among Vietnamese EFL university students?

RQ2: What are the effects of AI-mediated low-stakes speaking practice on students' oral performance as measured by an analytic speaking rubric?

RQ3: How do students in the experimental group perceive the role of AI-mediated practice in their speaking anxiety reduction and confidence development?

2. Literature Review

2.1. Methodological Evolution in AI-Mediated Speaking Research

The rapid proliferation of generative AI tools has prompted a wave of review studies mapping the evolving landscape. Crompton et al. (2024), in a systematic review of 42 studies on AI in English language teaching, found that speaking was among the skills most frequently supported by AI, yet affective outcomes remained underexamined. Koç and Savaş (2025), in a meta-synthesis of 57 studies on voice-based AI chatbots, noted a conspicuous absence of research applying the FLA framework to AI-mediated speaking practice. Earlier work focused primarily on cognitive and linguistic outcomes. Bibauw et al. (2019) synthesized dialogue-based CALL findings and reported moderate effects on vocabulary and grammar, noting that affective outcomes were rarely primary variables. Huang et al. (2022) concluded that most chatbot studies emphasized linguistic accuracy over anxiety. Kim et al. (2019) and Jeon (2024) provided early qualitative evidence that learners valued non-judgmental AI interaction, but both lacked control groups. The landscape has since improved: Fathi et al. (2024) employed a true experimental design, Zhang et al. (2024) included validated anxiety instruments, and Chen et al. (2025) demonstrated that longer interventions (eight weeks) may yield more robust affective outcomes. Wu and Li (2026), in a recent meta-analysis, reported a substantial overall

effect of AI-enhanced instruction on speaking proficiency, though affective outcomes were not separately meta-analyzed. However, no study has examined sub-component effects across the FLA construct (CA, FNE, TA separately). In summary, existing studies demonstrate promising effects of AI-mediated speaking practice, but three limitations persist: limited attention to affective mechanisms beyond total anxiety scores, underrepresentation of Vietnamese and other ASEAN learners, and insufficient analysis of how AI practice differentially affects FLA sub-components. The present study addresses all three gaps.

2.2. The Vietnamese EFL Context: Sociocultural Amplifiers

Bao (2013) argued that Confucian-heritage values create a deference-to-authority dynamic rendering student-initiated speech an implicit norm violation. Truong and Ho (2022), surveying 297 Vietnamese freshmen, found that all three FLA components contributed to speaking anxiety, with assessment-related concerns particularly salient, suggesting that Vietnamese examination-oriented culture may intensify FLA beyond what is observed in contexts with more formative evaluation traditions. This profile suggests that an AI-mediated intervention could be particularly effective in Vietnam, but also raises the question of whether technology anxiety may be more pronounced in developing infrastructure contexts. Ngo (2023) found generally positive perceptions of ChatGPT among Vietnamese university students, but did not address speaking skills or anxiety.

2.3. Desirable Difficulties and the Affective Filter: Convergent Frameworks

This study draws on two convergent frameworks. Krashen's (1982) affective filter operates at the input-processing level: anxiety blocks comprehensible input. The desirable difficulties framework (Bjork & Bjork, 2011) operates at the practice-design level: optimal learning requires challenge without threat. They converge in predicting that low-stakes AI practice should benefit learners but differ in mechanism—reduced input blockage (Krashen) versus optimized practice conditions (Bjork & Bjork). The mixed-methods design allows exploration of both: quantitative anxiety reduction maps onto the affective filter, while qualitative reports of automatization map onto desirable difficulties. Additionally, DeKeyser's (2007) skill acquisition framework predicts that repeated practice develops procedural knowledge, enabling more fluent production—directly testable through dimension-level speaking performance analysis.

3. Methodology

3.1. Research Design

This study employed a quasi-experimental pretest–posttest non-equivalent control group design (Creswell & Creswell, 2018). Two intact classes of second-year English-major students at a Vietnamese university were assigned as experimental (n = 32) and control (n = 32) groups. The two classes were purposively selected based on comparable prior-semester GPA in English skills courses, equivalent CEFR placement results (B1 or above, as determined by the university's standardized placement test), and similar instructor-reported participation patterns from the preceding semester.

3.2. Participants

A total of 64 second-year English-major students participated. All were Vietnamese nationals aged 19–21 at a public university in Ho Chi Minh City. The two intact classes each enrolled 43 students; however, participation was voluntary, and only students who provided informed consent were included. In each class, 32 students consented, while the remaining students—most of whom were retaking the course and declined the additional practice

commitment—were not enrolled in data collection, yielding a final sample of 64. The experimental group consisted of 32 students (24 female, 8 male) and the control group of 32 students (22 female, 10 male). All had achieved at least B1 on the CEFR as determined by the university's standardized placement test. Prior AI experience was assessed via a screening questionnaire administered during recruitment; no participant reported prior sustained use of conversational AI tools for English speaking practice. Ethical approval was obtained from the university's Institutional Review Board (IRB), and all participants provided written informed consent. No participants withdrew during the study period; all 64 completed both pretest and posttest assessments.

3.3. Instruments

3.3.1. Speaking anxiety questionnaire

Speaking anxiety was measured using a 20-item questionnaire adapted from the 33-item FLCAS (Horwitz et al., 1986). Items were selected through a two-stage process: first, two applied linguistics experts independently identified items specifically relevant to speaking anxiety (excluding items pertaining solely to reading, writing, or general classroom behavior); second, items with factor loadings below .40 in prior validation studies were excluded. The resulting 20-item instrument comprised communication apprehension (8 items), fear of negative evaluation (7 items), and test anxiety (5 items) subscales. Items used a five-point Likert scale (1 = strongly disagree to 5 = strongly agree; total range: 20–100). The instrument was piloted with 30 students from a non-participating class. Cronbach's alpha for the total scale was .89; subscale alphas were .84 (CA), .86 (FNE), and .78 (TA), all exceeding the .70 threshold recommended for research instruments. A Vietnamese translation was provided following back-translation procedures (Brislin, 1970).

3.3.2. Analytic speaking rubric

Oral performance was assessed using an analytic rubric adapted from Luoma (2004), evaluating fluency, pronunciation, grammatical accuracy, vocabulary range, and interactional competence on a 1–5 scale (maximum 25). Two trained raters—both experienced EFL lecturers—underwent a three-hour calibration session in which they independently scored 10 practice samples, discussed discrepancies, and reached consensus on scoring criteria. During data collection, both raters independently scored all 128 speaking samples (64 participants × 2 time points) while blind to group membership and time point. Scoring disagreements exceeding one point on any dimension were resolved through discussion to consensus. A two-way mixed-effects ICC computed on all 128 samples yielded .91, indicating excellent agreement (Koo & Li, 2016).

3.3.3. Semi-structured interview guide

A semi-structured interview guide consisting of eight open-ended questions addressed students' experiences with the AI tool, perceived changes in anxiety and confidence, and suggestions for improvement. Twelve experimental-group participants were selected based on pretest FLCAS scores: four from the top quartile (high anxiety), four from the middle quartiles (moderate), and four from the bottom quartile (low anxiety). Interviews were conducted individually in Vietnamese by a research assistant not involved in teaching the course, to minimize social desirability effects. Each interview lasted approximately 20–30 minutes, was audio-recorded with participant consent, and was transcribed verbatim. Transcripts were member-checked by sending summaries to participants for verification.

3.4. Intervention Design

The six-week intervention was embedded within the regular Speaking Skills II course. Both groups received the same weekly two-hour face-to-face class, taught by the same instructor using identical lesson plans and materials, to control for instructor effects. The experimental group additionally completed three 30-minute AI-mediated speaking sessions per week using ChatGPT (GPT-4o voice mode), which supports real-time voice-based conversation and automatic speech recognition. The platform provides implicit corrective feedback through conversational recasts, pronunciation modeling via text-to-speech output, and vocabulary suggestions within conversational context, though it does not offer systematic grammar correction or scored pronunciation assessment. Compliance was monitored through weekly self-report logs verified against conversation screenshots (mean compliance: 87%; range: 72–97%). Activities were scaffolded:

Weeks 1–2: Guided role-plays. Structured scenarios with scripted prompts and vocabulary support.

Weeks 3–4: Semi-guided conversations. Topic-based conversations with minimal structural support.

Weeks 5–6: Open-ended discussions. Self-selected topics simulating naturalistic spoken interaction.

The control group completed equivalent speaking preparation tasks, including recorded self-practice monologues, textbook-based speaking prompts, and vocabulary rehearsal logs (approximately 90 minutes/week), but without interactive AI-mediated oral practice. Thus, the comparison is between AI-mediated interactive speaking practice and non-interactive conventional preparation tasks.

3.5. Data Collection

At both pretest (Week 1) and posttest (Week 6), data collection followed a fixed sequence: participants first completed the speaking anxiety questionnaire in a private setting, then proceeded to the individual speaking assessment. This order was chosen to capture anxiety levels prior to the potential stress of the speaking task. Speaking assessments involved a three-minute conversation with a trained interlocutor (the same interlocutor at both time points) on a topic drawn randomly from a pool of 10 pre-validated topics. Different topics were assigned at pretest and posttest to minimize practice effects. Conversations were audio-recorded and rated by two raters blind to group membership and time point.

3.6. Data Analysis

Quantitative data were analyzed using SPSS (Version 28). Mixed-design ANOVAs examined changes in anxiety and performance. Effect sizes were reported using η^2p and Cohen's d (calculated as the within-group mean difference divided by the pooled standard deviation of pretest and posttest scores) with 95% confidence intervals. An ANCOVA was performed with pretest scores as covariate. Normality (Shapiro-Wilk) and homogeneity of variance (Levene's) were verified. Given the multiple sub-component (3) and dimension-level (5) analyses, a Bonferroni-adjusted alpha of .006 was applied for the sub-component tests and .01 for the dimension-level tests; however, results are also reported at the conventional .05 level for transparency. Qualitative data were analyzed inductively using thematic analysis (Braun & Clarke, 2006): transcripts were open-coded without a priori categories, and codes were iteratively grouped into themes. A second researcher independently coded 25% of transcripts ($\kappa = .81$).

3.7. Delimitations

This study is delimited in several respects. First, it is confined to second-year English-major students at a single public university in Ho Chi Minh City; findings may not generalize to non-English majors, private universities, or other regions. Second, the AI tool is limited to ChatGPT (GPT-4o voice mode); other platforms may yield different results. Third, the study examines short-term effects over six weeks; long-term retention falls outside its scope. Fourth, speaking performance is assessed via a researcher-designed analytic rubric rather than a standardized proficiency test. Fifth, the study does not measure technology proficiency or digital literacy as moderating variables.

4. Results

4.1. Baseline Comparability

Independent samples t-tests confirmed no significant pretest differences for total speaking anxiety, $t(62) = 0.21$, $p = .834$, $d = 0.05$, or speaking performance, $t(62) = 0.23$, $p = .820$, $d = 0.06$, nor across any anxiety sub-component (all $p > .70$).

4.2. Effects on Speaking Anxiety (RQ1)

Normality was confirmed (Shapiro-Wilk, all $p > .05$) and homogeneity of variance satisfied (Levene's, all $p > .70$). Descriptive statistics are presented in Table 1.

Table 1

Descriptive Statistics for Total Speaking Anxiety Scores by Group and Time

Group	Pre M	Pre SD	Post M	Post SD	Diff.	d [95% CI]
Exp. (n=32)	72.41	9.83	58.63	10.14	−13.78	1.38 [0.98, 1.78]
Ctrl. (n=32)	71.88	10.27	69.22	9.91	−2.66	0.26 [−0.13, 0.65]

The mixed-design ANOVA revealed a significant main effect of time, $F(1, 62) = 48.73$, $p < .001$, $\eta^2p = .44$; a significant Group × Time interaction, $F(1, 62) = 31.56$, $p < .001$, $\eta^2p = .34$; and a significant main effect of group, $F(1, 62) = 7.41$, $p = .008$, $\eta^2p = .11$. Post hoc pairwise comparisons indicated that the experimental group's anxiety scores decreased significantly from pretest to posttest ($M = 72.41$ to $M = 58.63$; mean difference = 13.78, $p < .001$, $d = 1.38$, 95% CI [0.98, 1.78]), while the control group showed a small, non-significant decrease ($M = 71.88$ to $M = 69.22$; mean difference = 2.66, $p = .187$, $d = 0.26$, 95% CI [−0.13, 0.65]). The ANCOVA with pretest scores as covariate confirmed these findings, $F(1, 61) = 29.84$, $p < .001$, $\eta^2p = .33$, indicating that the between-group difference was not attributable to pretest variation.

Sub-component analysis (Table 2) revealed differential effects. The experimental group showed reductions across all three sub-components: communication apprehension decreased by 6.22 points ($d = 1.29$), fear of negative evaluation by 5.94 points ($d = 1.47$), and test anxiety by 1.62 points ($d = 0.44$). In standardized terms, the effect was largest for FNE, followed by CA, with TA showing only a small effect. Significant Group × Time interactions were found for FNE, $F(1, 62) = 24.81$, $p < .001$, $\eta^2p = .29$, and CA, $F(1, 62) = 18.63$, $p < .001$, $\eta^2p = .23$, but not for TA, $F(1, 62) = 1.94$, $p = .169$, $\eta^2p = .03$. This pattern confirms the theoretical

prediction that AI-mediated low-stakes practice primarily targets evaluative and communicative anxiety while leaving assessment-related anxiety relatively unchanged.

Table 2

Anxiety Sub-Component Scores by Group and Time

Component	*Group*	*M*	*SD*	*M*	*SD*	*Diff.*	*d*
CA (8 items)	Exp.	29.38	4.71	23.16	4.93	−6.22	1.29
	Ctrl.	28.94	5.02	27.88	4.86	−1.06	0.21
FNE (7 items)	Exp.	25.47	3.88	19.53	4.22	−5.94	1.47
	Ctrl.	25.19	4.13	24.13	3.97	−1.06	0.26
TA (5 items)	Exp.	17.56	3.41	15.94	3.87	−1.62	0.44
	Ctrl.	17.75	3.53	17.21	3.44	−0.54	0.15

*CA = Communication Apprehension; FNE = Fear of Negative Evaluation; TA = Test Anxiety.

4.3. Effects on Speaking Performance (RQ2)

The mixed-design ANOVA revealed a significant Group × Time interaction, $F(1, 62) = 12.37$, $p < .001$, $\eta^2p = .17$ (Table 3). The experimental group's mean performance score increased from 14.72 (SD = 3.18) to 18.41 (SD = 2.96), a gain of 3.69 points (d = 1.20, 95% CI [0.82, 1.58]). The control group also improved, from 14.53 (SD = 3.41) to 15.88 (SD = 3.27), a gain of 1.35 points (d = 0.40, 95% CI [0.02, 0.78]). Dimension-level analysis (Table 4) indicated significant Group × Time interactions for fluency ($p < .01$), vocabulary ($p < .05$), and interactional competence ($p < .01$), but non-significant interactions for pronunciation ($p = .083$) and grammatical accuracy ($p = .114$).

Table 3

Total Speaking Performance Scores by Group and Time

	Pretest		*Posttest*			
Group	*M*	*SD*	*M*	*SD*	*Diff.*	*d [95% CI]*
Exp. (n=32)	14.72	3.18	18.41	2.96	+3.69	1.20 [0.82, 1.58]
Ctrl. (n=32)	14.53	3.41	15.88	3.27	+1.35	0.40 [0.02, 0.78]

Table 4

Speaking Performance by Rubric Dimension

Experimental			Control			
Dimension	*Pre*	*Post*	*Gain*	*Pre*	*Post*	*Gain*	*p*
Fluency	2.72	3.78	1.06	2.69	3.03	0.34	<.01
Pronunc.	2.91	3.44	0.53	2.88	3.13	0.25	.083
Grammar	3.03	3.50	0.47	2.97	3.22	0.25	.114
Vocabulary	3.12	3.81	0.69	3.05	3.28	0.23	<.05
Interact.	2.94	3.88	0.94	2.94	3.22	0.28	<.01
Total	**14.72**	**18.41**	**3.69**	**14.53**	**15.88**	**1.35**	**<.001**

4.4. Student Perceptions (RQ3)

Thematic analysis of the 12 interviews yielded four themes addressing the two dimensions of RQ3: Themes 1 and 2 illuminate the anxiety-reduction dimension, Theme 3 addresses the confidence-development dimension, and Theme 4 captures perceived limitations that contextualize the overall experience.

4.4.1. Reduced fear of judgment (10/12)

Ten participants identified the absence of social evaluation as the primary benefit. Hoa (high-anxiety) explained that with the AI, she could focus entirely on expressing ideas. Minh (moderate-anxiety) noted he could attempt longer sentences without fear of losing face. Two participants (Linh and Quang) did not identify this as primary, noting they had not experienced high evaluative anxiety prior to the intervention.

4.4.2. Benefits of repetition and self-pacing (9/12)

Nine participants valued repeating scenarios. Trang practiced a restaurant role-play seven times; by the fourth attempt, she no longer paused for vocabulary. Duc appreciated the AI's patience. Three participants found repetition less valuable, with An noting it became monotonous after the third week.

4.4.3. Gradual confidence building (8/12)

Eight participants described increasing confidence that transferred to classroom settings. Lan, who had never voluntarily spoken in class during three semesters, raised her hand after four weeks of AI practice. Four participants did not report classroom transfer, though they acknowledged greater comfort with the AI itself.

4.4.4. Perceived limitations (12/12)

All 12 identified limitations. Nine mentioned speech recognition errors with Vietnamese-accented English. Seven noted limited conversational depth. Five expressed that the AI could not replace human interaction's spontaneity. Duc captured the prevailing sentiment: the AI was like a gym—useful for building strength but not the same as playing a real sport.

5. Discussion

The central finding is straightforward: six weeks of AI-mediated practice produced a large reduction in overall speaking anxiety. The experimental group demonstrated a large anxiety reduction ($d = 1.38$, 95% CI [0.98, 1.78]), consistent with findings from Iranian (Fathi et al., 2024), Chinese (Chen et al., 2025; Zhang et al., 2024), and Indonesian (Wu et al., 2026) contexts. This study extends the evidence to Vietnamese higher education.

What sets this study apart, however, is the sub-component analysis. FNE showed the largest reduction ($d = 1.47$), CA the second ($d = 1.29$), while the TA interaction was non-significant ($p = .169$). This differential pattern—unprecedented in the AI-mediated speaking literature, where previous studies report only total scores—aligns with the theoretical rationale: AI removes the evaluative audience (reducing FNE) and provides private interaction (reducing CA) but does not simulate formal assessment (leaving TA unchanged). The qualitative findings directly support this quantitative pattern. Theme 1 (reduced fear of judgment), reported by 10 of 12 participants, corresponds to the FNE reduction; Theme 2 (benefits of repetition and self-pacing) corresponds to the CA reduction, as private practice removes the real-time audience

pressure that triggers communication apprehension. That TA was minimally affected is consistent with Theme 4, in which no participant reported reduced test-related concerns.

The performance data tell a complementary story. Fluency and interactional competence showed the largest gains, consistent with Wu et al. (2026). Pronunciation ($p = .083$) and grammatical accuracy ($p = .114$) did not show significant differential improvement, indicating that conversational AI practice alone is insufficient for accuracy-related skills requiring explicit metalinguistic instruction. These limitations mirror the system constraints noted by Hsu et al. (2023) and El Shazly (2021). The practical takeaway is clear: AI practice should be supplemented with targeted pronunciation and grammar instruction. The control group also improved modestly ($d = 0.40$, 95% CI [0.02, 0.78]), likely attributable to the regular classroom instruction both groups received. The differential gain, however, demonstrates that AI-mediated practice conferred benefits substantially beyond what conventional instruction alone provided.

The qualitative finding that participants developed automatic responses through repetition is consistent with DeKeyser's (2007) skill acquisition framework: procedural knowledge develops through practice, enabling production with reduced cognitive effort. The transfer of confidence to classroom settings, reported most vividly by Lan, suggests that automaticity gained through AI practice frees cognitive resources during face-to-face interaction. Duc's gym metaphor captures the complementary role participants attributed to AI practice, a view consistent with sociocultural theory's insistence that authentic human interaction remains essential for full communicative development (Lantolf & Thorne, 2006).

Lan's trajectory is not an isolated case but rather representative of a broader pattern observable in Vietnamese EFL classrooms: once a learner breaks the initial barrier of voluntary speaking — even once — subsequent participation becomes progressively easier. What the AI practice appears to have provided is a safe rehearsal space where that first breakthrough could occur without the social risk that the classroom entails. Various strategies have been attempted to reduce speaking anxiety in this context — pair work, anonymous response systems, small-group rotations — but none addresses the root problem as directly as removing the human audience altogether. The AI tool does precisely this, and the fact that the confidence built in private practice transferred to public classroom participation suggests that the underlying mechanism is not avoidance of speaking but rather graduated exposure. This raises the practical question of how to systematically scaffold learners from low-stakes AI practice toward high-stakes human interaction. The present findings suggest a graduated transfer pathway: after the initial weeks of private AI practice, instructors can introduce progressively more exposed speaking contexts—first paired peer conversations, then small-group discussions, and finally whole-class presentations—so that the confidence built in the non-evaluative AI environment is incrementally tested against increasingly authentic social conditions. In this model, AI practice is not an endpoint but the first rung of an anxiety-graded ladder, with each subsequent step reintroducing a controlled degree of the evaluative pressure that AI practice initially removes.

5.1. Limitations and Directions for Future Research

Several limitations should be acknowledged. First, the quasi-experimental design does not permit full causal inference. Second, the novelty of AI interaction may have generated a Hawthorne effect: engagement with novel technology, rather than the specific anxiety-reducing affordances, may partially explain results. Future research should include a technology-active control group (e.g., using AI for non-speaking tasks) to isolate oral AI practice effects from technology novelty. Third, despite compliance monitoring (87%), self-reported logs may

overestimate engagement. Fourth, the sample was limited to English-major students at a single university, and the gender distribution (predominantly female in both groups) may limit generalizability given documented gender differences in FLA (Teimouri et al., 2019). Fifth, the six-week period may not capture long-term trajectories or regression effects. A further consideration concerns technological equity: although all participants in this study had reliable access to a smartphone and internet connection, this cannot be assumed universally across Vietnamese higher education, where disparities in device ownership and network quality persist between urban and rural student populations. Instructors adopting AI-mediated practice should audit student access in advance and provide institutional alternatives—such as computer-lab sessions or shared devices—to ensure the intervention does not inadvertently disadvantage learners with limited technological resources. Future studies should employ physiological anxiety measures, delayed posttests, multi-site designs, and comparisons of AI platform types (cf. Wang et al., 2024).

6. Conclusion

This study makes two primary contributions. First, it is the first empirical study to examine AI-mediated speaking practice and language anxiety in Vietnamese higher education, extending the evidence base beyond the East Asian, Middle Eastern, and Southeast Asian contexts investigated to date. Second, it is the first AI intervention study to demonstrate differential effects across FLA sub-components, providing granular support for the theoretical claim that AI practice targets evaluative and communicative anxiety rather than producing undifferentiated reduction. Practically, the findings offer a concrete strategy for integrating AI tools as a structured supplement—not a replacement—for conventional speaking instruction. In the context of Vietnamese higher education, where speaking anxiety remains persistent and culturally reinforced, AI-mediated low-stakes practice may serve as a promising supplementary strategy, particularly when integrated with teacher-guided speaking instruction and targeted pronunciation and grammar practice. It should be emphasized that these findings do not suggest replacing teacher-led speaking instruction with AI tools. Rather, they point to a specific and limited role for AI: as a confidence-building bridge between the safety of silent knowledge and the risk of public speech. In the Vietnamese university context, where class sizes make individualized speaking practice virtually impossible and where cultural norms discourage the trial-and-error that fluency development requires, this bridge is not a luxury but a pedagogical necessity. These findings translate into several actionable recommendations for EFL instructors. First, conversational AI should be introduced as a structured homework component—three 30-minute sessions per week with scaffolded tasks progressing from guided role-plays to open-ended discussions over a six-week cycle. Second, AI practice should be explicitly positioned as a fluency and interaction builder; instructors should continue providing direct pronunciation instruction and explicit grammar teaching in class. Third, a one-week familiarization phase should precede formal practice, as qualitative data indicate technology-related frustrations diminish rapidly after initial sessions. Fourth, compliance should be monitored through weekly logs with screenshot verification. Fifth, periodic debriefing sessions should be conducted in which students discuss their AI practice experiences, facilitating metacognitive awareness of anxiety reduction. These guidelines are implementable within existing Vietnamese university curricula without requiring additional institutional resources beyond student access to a smartphone or computer with internet connectivity. Crucially, this design integrates into existing curricula without encroaching on regular class hours: the AI sessions function as structured, self-paced homework completed outside class, while in-class time is reserved for teacher-led feedback, error correction, and the graduated peer-interaction activities described above. This arrangement follows a flipped-classroom logic, in which low-

stakes rehearsal occurs privately at home and valuable face-to-face class time is devoted to the higher-order communicative work that AI cannot replicate. For a typical three-credit speaking course, the three weekly AI sessions add roughly ninety minutes of out-of-class practice—comparable to conventional homework loads—and therefore impose no additional burden on scheduled instruction.

Notes

1 Corresponding author, Lecturer, Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam; Email: leminhhang@tdtu.edu.vn

References

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal, 78*(2), 155–168. https://doi.org/10.1111/j.1540-4781.1994.tb02026.x

Bao, D. (2013). Voices of the reticent? Getting inside views of Vietnamese secondary students on learning. In M. Cortazzi & L. Jin (Eds.), *Researching cultures of learning: International perspectives on language learning and education* (pp. 136–154). Palgrave Macmillan. https://doi.org/10.1057/9781137296344_8

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning, 32*(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Chen, Y., Ke, N., Huang, L., & Luo, R. (2025). The role of GenAI in EFL speaking: Effects on oral proficiency, anxiety and risk-taking. *RELC Journal*. Advance online publication. https://doi.org/10.1177/00336882251341072

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology, 55*(6), 2503–2529. https://doi.org/10.1111/bjet.13460

DeKeyser, R. M. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667275

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), Article e12667. https://doi.org/10.1111/exsy.12667

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121*, Article 103254. https://doi.org/10.1016/j.system.2024.103254

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior, 93*, 279–289. https://doi.org/10.1016/j.chb.2018.12.023

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hsu, M.-H., Chen, P.-S., & Yu, C.-S. (2023). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments, 31*(7), 4297–4308. https://doi.org/10.1080/10494820.2021.1960864

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257. https://doi.org/10.1111/jcal.12610

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning, 37*(1–2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

Kim, N.-Y., Cha, Y., & Kim, H.-S. (2019). Future English learning: Chatbots and artificial intelligence. *Multimedia-Assisted Language Learning, 22*(3), 32–53. https://doi.org/10.15702/mall.2019.22.3.32

Koç, F. Ş., & Savaş, P. (2025). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL, 37*(1), 4–21. https://doi.org/10.1017/S0958344024000168

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71–86. https://doi.org/10.1111/j.1540-4781.2008.00687.x

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press. https://doi.org/10.1017/CBO9780511733017

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

McLaughlin, B. (1987). *Theories of second-language learning*. Edward Arnold.

Ngo, T. T. A. (2023). The perception by university students of the use of ChatGPT in education. *International Journal of Emerging Technologies in Learning, 18*(17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition, 41*(2), 363–387. https://doi.org/10.1017/S0272263118000311

Tran, T. T. T., & Moni, K. (2015). Management of foreign language anxiety: Insiders' awareness and experiences. *Cogent Education, 2*(1), Article 992593. https://doi.org/10.1080/2331186X.2014.992593

Truong, T. N. N., & Ho, D. P. K. (2022). Exploring Vietnamese non-English-majored freshmen's English-speaking anxiety at a public university in Vietnam. *VNU Journal of Foreign Studies, 38*(5), 105–125. https://doi.org/10.25073/2525-2445/vnufs.4789

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System, 127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal, 37*(3), 308–328. https://doi.org/10.1177/0033688206071315

Wu, T.-T., Hapsari, I. P., & Huang, Y.-M. (2026). Effects of incorporating AI chatbots into think–pair–share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning, 39*(5–6), 1408–1446. https://doi.org/10.1080/09588221.2025.2478271

Wu, X., & Li, R. (2026). Unraveling effects of AI chatbots on EFL learners' language skill development: A meta-analysis. *The Asia-Pacific Education Researcher, 35*(4), 703–714. https://doi.org/10.1007/s40299-024-00853-2

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal, 75*(4), 426–439. https://doi.org/10.1111/j.1540-4781.1991.tb05378.x

Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System, 121*, Article 103259. https://doi.org/10.1016/j.system.2024.103259

Bionote

Le Minh Hang is a lecturer at the Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam. She holds an MA in TESOL from the University of Huddersfield, United Kingdom. Her research interests include English language teaching methodology, AI literacy, language skills development, and technology-enhanced language learning. Her current work focuses on the integration of artificial intelligence into English language teaching and teacher professional development. Email: leminhhang@tdtu.edu.vn

#80 INSTRUCTORS' DIGITAL COMPETENCE AFFECTING STUDENT FEEDBACK AND PERSONALIZED LEARNING IN UNIVERSITY EDUCATION

Nguyen Thanh Tuan[1], Nguyen Phuong Bao Tran[2]

Abstract

In the context of digital transformation, instructors' digital competence has become a crucial factor affecting teaching quality and students' learning experience. Digital competence is not simply a capability of using technology, but it also includes integrating digital tools to design lessons, evaluate, provide feedback, and support learning effectively. However, feedback practice and personalized learning are still a research gap, especially in university education; both lecturers and students need to adapt to digital technology. The study aims to discover the effect of digital competence on students' feedback and personalized learning. The quantitative approach with a cross-sectional survey design is mainly applied to this study, including observation, document analysis, and a questionnaire. Data collection is done through the questionnaire, measuring three main variables: instructors' digital competence, feedback quality, and personalized learning level. Especially, the data analysis is based on the SEM model (SmartPLS4) to test research hypotheses and evaluate the relationship among variables. The expected results show the positive effect of digital competence on feedback effectiveness and personalized learning. Instructors with good capabilities of using technology can integrate digital tools to provide in-time feedback, transparency, and flexibility to meet learning needs. Effective feedback, which can strongly promote autonomous learning by supporting students with self-regulated strategies, is the intermediary element between digital competence and personalized learning. The study significantly contributes to developing professional training programs for instructors, establishing a theoretical foundation of digital competence, creating a flexible learning environment with a student-centered approach, and providing positive feedback to meet students' learning needs effectively.

Keywords: digital competence, student feedback, personalized learning, self-regulated learning, autonomous learning, digital technology, online teaching…

1. Introduction

The rapid change of digital competence has influenced teaching and learning approaches in university education in Viet Nam, Particularly Hong Bang International University. In this context, instructors' digital competence is not only about the use of technology but also about the ability to integrate it effectively into teaching activities. One of the factors directly affected by instructors is feedback, considered the most powerful tool for improving students' learning outcomes. However, the effectiveness of feedback depends on how it is designed, personalized, and how HIU instructors and students interact. Alongside that, personalized learning is increasingly receiving attention as a modern educational trend, in which content and learning approach need to be adjusted to fit each student. Digital technology plays an important role in personalized support, but this belongs to instructors' digital

[1] *Can Tho University; Hong Bang International University*
[2] *Can Tho University*

competence. Although there have been numerous studies on individual factors, the relationships among instructors' digital competence, feedback, and personalized learning haven't been fully explored. Based on the research gap identified, this study was conducted to answer the following questions: (RQ1) How does instructors' digital competence affect the effectiveness provided to students? (RQ2) How does the instructors' digital competence affect the implementation of personalized learning in university education? (RQ3) Does the students' feedback play a mediating role in the relationship between lecturers' digital competence and personalized learning?

1.1 Problem Statement

Digital revolution has made a remarkable transformation in higher education. The previous studies have focused on the ability of using information technology since 2000s. However, the development of digital learning environment gradually changes not only using technology but also integrating it into teaching. In 2017, Redecker and Punie' study marked the significant alteration of digital competence. From 2016 to 2022, digital transformation boomed at high rate, and teaching online became a global issue. This became a decisive factor in teaching quality and learning experience. Hizam et al. (2021) stated the positive effect on the online learning system. In recent years, the studies have confirmed that digital competence have helped instructors access digital resources and innovate teaching approaches (Dang et al., 2024). Besides, student feedback is also considered an important part of online learning. Especially, digital feedback is a useful tool, helping students exchange information in time. Betancur-Chicué and García-Valcárcel Muñoz-Repiso (2023) certified that digital competence involves evaluation and feedback, one of limited factors of instructors at university currently. Artificial intelligence (AI) is a favorable condition for personalized learning development. The components of digital competence such as, digital resources, assessment, and learning support significantly contribute to student learning experience (Luketić & Diković, 2026).

1.2 Research Objective

The study aims to evaluate the role of digital competence in improving student learning experience through digital resources and the ability to use technology. It is essential to understand how instructors' digital competence affects online learning. Next, the study aims to analyze the influence of digital competence on student feedback and personalized learning. Finally, the study aims to support universities in Vietnam to develop digital competence programs, contributing to online teaching improvement and student-centered approach.

1.3 Research Scope

The study only focuses on surveying the relationship between instructors' digital competence and student feedback and personalized learning. The participants are instructors and students at HIU. The content considers three main variables: digital competence independent variable, students feedback and personalized learning (dependent variables), representing educational results affect by instructors' digital competence. The study does not mention other factors, such as learning motivation, learning outcomes, student satisfaction, and institutional issues. The research method applied quantitative approach through the survey. Data were collected in the definite time, and reflected the participants' perception at the timepoint of the survey.

2. Theoretical Foundation

2.1 The Concept of Instructors' Digital Competence

In the context of digital transformation in university education, Digital Competence (DC) is a core factor that determines teaching and learning effectiveness. This concept may be defined in various ways depending on the approach. However, most of the authors agree that DC is not only the ability to use technology but also includes cognitive, pedagogical, and social competencies in a digital environment. It is a dimensional concept, including the capability to use technology, digital knowledge, and the ability of technology in various learning contexts (Ilomäki et al., 2016). This shows that DC is integrative, and it can't be separated into technology skills and pedagogical competence.

In another study, it is the ability to design, implement, and evaluate teaching and learning activities in a digital environment. The DigCompEdu (Digital Competence Framework for Educators) is an international standardized framework for measuring instructors' digital competence, including competence dimensions: (1) professional engagement; (2) digital resources; (3) teaching and learning; (4) evaluation; (5) empowering learners; and (6) facilitating learners' digital competence. Especially, it emphasizes the role of instructors in using technology to personalize learning and provide students with effective feedback (Redecker, 2017). This proves that DC is an essential condition for promoting pedagogical innovation in university education.

Besides, other studies have confirmed that instructors' DC has a close relationship with the effectiveness of online teaching and student engagement. Instefjord and Munthe (2017) stated that instructors with high digital competence could integrate technology into teaching flexibly and creatively, thereby improving student learning experience. However, DC is not a static characteristic, but it is a process of ongoing development, affected by factors such as teaching experience, professional training, and institutional environment (Tondeur et al., 2017).

2.2 Student Feedback in Digital Learning

Feedback is one of the most important factors influencing the students' learning process. Hattie and Timperley (2007) emphasized that feedback could be defined as the information provided to students about their performance to close the gap between the current state and the desired learning goals. Feedback became very flexible in digital classes with technological support. Online learning platforms enabled instructors to provide immediate feedback and personalize each student's experience. This contributed to learning effectiveness and engagement (Nicol & Macfarlane-Dick, 2006). Feedback could be classified into several types: automated, instructor, and peer feedback. Personalized feedback was the strongest influence on students' learning outcomes. Effective feedback was needed to ensure the characteristics such as timeliness, specificity, guidance-oriented content, and suitability to students' needs (Shute, 2008). In a digital environment, Artificial Intelligence (AI) could support accurate feedback and fit each student. Besides, Carless and Boud (2018) added that it was not only a one-way transmission of information, but also an interactive process, and students played an active role in receiving and using feedback to improve their learning. It was important in digital classes, where students were required to have high self-regulation. However, providing students with feedback encountered challenges, including instructors' workload, insufficient digital skills, and limited technology use (Henderson et al.,2019). Therefore, instructors' DC plays an important role in ensuring feedback quality.

2.3 Personalized Learning

According to Pane et al. (2017), Personalized Learning (PL) is one of the outstanding trends in modern education. It is a content-adjusting method, pace, and learning approach to fit

the needs and the capabilities of each student. Also, it is supported by an adaptive learning system, AI, and data analysis. These tools enable instructors to collect and analyze data about students' learning behaviors, thereby offering content and appropriate learning activities (Holmes et al., 2019). As evidence, Walkington (2013) showed that PL could notably improve students' learning outcomes, motivation, and engagement. PL content according to students' preferences could enhance motivation and learning effectiveness. In addition, PL contributed to students' self-study and self-regulation competence, helping them become more proactive and independent learners (Zimmerman, 2002). This is significant for long-life learning. However, PL implementation faces many challenges, such as high technological requirements, data, and instructors' technological capabilities (Holmes et al., 2019). Therefore, their digital competence plays a core role in designing and performing PL activities.

2.4 The Relationship among DC, Feedback, and PL

DC is a foundational factor influencing feedback and PL in university education. Once instructors possess high digital competence, they can use digital tools effectively to provide immediate feedback and design appropriate learning activities with each student. Redecker (2017) confirmed that one of the important elements of digital competence was the capability of using technology to support and evaluate feedback. This indicated the direct relationship between DC and feedback quality in digital classes. Additionally, DC enabled instructors to explore learning data to understand students' needs and characteristics, thereby designing PL activities (Holmes et al., 2019). This contributes to enhancing learning outcomes and students' satisfaction. Some empirical studies have argued that instructors' DC positively affects students' engagement and learning outcomes through improving the quality of feedback and PL (Tondeur et al., 2017). However, this relationship is not always direct, but it can be affected by the mediating factors, such as instructors' attitudes toward technology, institutional support, and infrastructure (Instefjord & Munthe, 2017).

2.5 Research Gaps

Although many studies about DC, feedback, and PL have been explored, some notable gaps still remain. First, most studies focus on students' DC rather than on instructors', while the role of instructors is very pivotal (Tondeur et al., 2017). Second, the studies have explored the factors in isolation rather than building an integrated model to analyze the relationship among DC, feedback, and PL. Finally, the studies are still limited in the Vietnamese setting, leading to a lack of appropriate empirical evidence for the local context. Therefore, this study addresses the gaps by building an integrated model to analyze instructors' DC affecting student feedback and PL in university education.

2.6 Hypotheses

From the above theoretical basis and research gaps, four research hypotheses were proposed, including three direct effects (H1-H3) and a mediating effect (H4).

H1. Instructors' digital competence has a positive effect on student feedback.

H2. Instructors' digital competence has a positive effect on personalized learning.

H3. Student feedback has a positive effect on personalized learning.

H4. Student feedback plays a mediating role between digital competence and personalized learning.

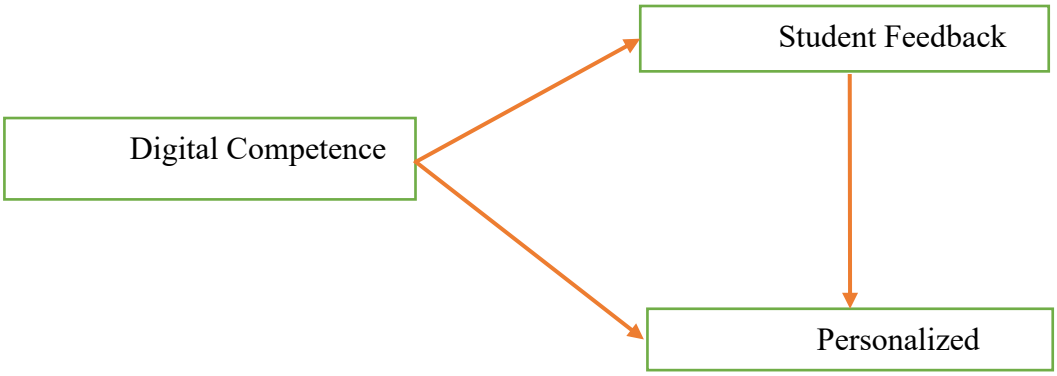
3. Methodology

3.1 Research Design

This study used a quantitative approach to test hypotheses, involving the influence of instructors' DC on feedback and PL in university education. The quantitative approach was selected to measure the research variables as quantitative data and to test causal relationships between variables using statistical techniques (Creswell & Creswell, 2018). In addition, the explanatory research design was also applied to the study. The data were collected via a questionnaire and analyzed using PLS-SEM. This was quite suitable for the studies whose objectives are to test theoretical models and the relationships between latent variables (Hair et al., 2022).

3.2 Research Model

On the theoretical foundation of DigCompEdu (Redecker, 2017), feedback (Hattie & Timperley, 2007), and personalized learning (Pane et al., 2017), the study proposed a three-variable research model, including instructors' DC, student feedback, and PL.



The proposed research model indicates that instructors' digital competence serves as a foundation for improving the quality of feedback and supporting personalized learning. For example, instructors with high digital competence can use technology to provide in-time and appropriate feedback to help students to self-regulate the learning process according to individual needs. At the same time, DC has a direct effect on implementing PL activities without feedback. Besides, students' feedback is also an important mediating mechanism to transform instructors' DC into PL experience. This is quite appropriate with the previous studies, proving that effective feedback is the core factor for improving learning outcomes and student engagement, and instructors' DC can foster interactive quality and technical support (Tondeur et al., 2017).

3.3 Subjects and Sampling Size

The survey participants were students at Hong Bang International University (HIU), enrolling in technology-enhanced courses, including online learning and blended learning. The reasons for selecting this subject were that they directly experienced instructors' DC, received feedback, and engaged in PL activities. Therefore, this was a source of appropriate data to evaluate perceived variables (Hattie & Timperley, 2007). The sampling selection is convenient to ensure the feasibility of data collection. According to Hair et al. (2022), for the PLS-SEM model, the minimum sample size ranged from 150 to 200 observations, increasing reliability. However, the study collected 246 samples to ensure validation.

3.4 Data Collection Instruments

The data were collected through the questionnaire using a 5-point Likert scale, ranging from levels: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree. The scales were based on the previous studies to ensure reliability and validity. Particularly, those were the scale of instructors' DC (Redecker, 2017), the scale of student feedback inherited from earning feedback (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006), and the scale of personalized learning (Pane et al., 2017; Holmes et al., 2019). The use of scales was tested in the previous studies; therefore, they increased the reliability and validity of data collection instruments (Hair et al., 2022).

3.5 Procedure for Data Collection and Analysis

The procedure for collecting data was performed in the following steps. First, the questionnaire was designed based on the scales of previous studies. Second, the pilot test was done to test the transparency and understandability of questions. Based on the results of the pilot test, the questionnaire was adjusted before the main survey. The main data was collected through Google Forms in a defined period. Finally, the data were cleansed to eliminate the irrelevant questions before the analysis performance. This ensured the reliability of data collection (Creswell & Creswell, 2018).

Data was analyzed through SmartPLS with PLS-SEM. The analysis process consisted of measurement assessment and structural model assessment. First, the measurement assessment was tested through the coefficient of Cronbach's Alpha (≥ 0.7), the outer loading (≥ 0.7), Composite Reliability (Rh-c ≥ 0.7), and Average Variance Extracted (AVE ≥ 0.5) to evaluate reliability and convergence. Besides, the discriminant validity was assessed using the Heterotrait–Monotrait ratio (HTMT) index (<0.85) (Hair et al., 2022). Second, the structural model assessment was implemented through the path coefficients and bootstrap to test the statistical significance of hypotheses (Hair et al., 2022).

4. Result Analysis & Discussion

Measurement model assessment and structural model assessment on SmartPLS aim to test reliability and validity. The analysis results are presented in the tables below. These results determine the appropriateness of the model as a solid foundation to discuss the relationships among variables.

Table 1

Outer loadings – Matrix

	DC	PL	SF
DC1	0.857		
DC2	0.859		
DC3	0.847		
DC4	0.822		
DC5	0.841		
PL1		0.880	
PL2		0.865	
PL3		0.904	

PL4		0.907	
PL5		0.887	
SF1			0.861
SF2			0.859
SF3			0.881
SF4			0.863
SF5			0.901

Table 1 indicates the outer loadings of the indicators for three constructs: Digital Competence (DC), Personalized Learning (PL), and Student Feedback (SF). The results illustrate that all indicator loadings exceed the recommended threshold of 0.70. ranging from 0.822 to 0.907. Particularly, the DC construct ranges from 0.822 to 0.859, the PL construct from 0.865 to 0.907, and the SF construct from 0.859 to 0.901. These loadings demonstrate a strong relationship with respective constructs and adequately represent the latent variables. As a result, the measurement model has achieved satisfactory indicator reliability and convergence validity.

Table 2

Construct Reliability and Validity - Overview

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
DC	0.900	0.903	0.926	0.715
PL	0.933	0.935	0.949	0.790
SF	0.922	0.923	0.941	0.762

Table 2 shows the convergent validity of latent variables, including DC, PL, and SF. The value of Cronbach’s alpha was larger than 0.70, fluctuating from 0.900 to 0.933. This reflected the high reliability of scales. Similarly, CR was also over 0.70, reflecting the consistency and stability of scales. Besides, EVA was from 0.715 to 0.790, exceeding 0.50. It was obvious that the observed variables adequately explained the variance of their corresponding latent constructs and satisfied the requirement for convergent validity. This result confirmed that the measurement model reached reliability and validation for the next analysis.

Table 3

Discriminant validity -HTMT Matrix

	DC	PL	SF
DC			

PL	0.841		
SF	0.837	0.835	

Table 3 shows the discriminant validity of concepts of latent variables, including DC, PL, and SF. The HTMT coefficient between PL and DC was 0.861, SF and DC (0.867), and SF and PL (0.875).

Table 4

R-Square Overview

	R-square	R-square adjusted
PL	0.852	0.851
SF	0.783	0.782

Table 4 presents the explanation of the model through R-square overview. The R-square value of PL was 0.852, and the R-square adjusted was 0.851, while the R-square of SF was 0.783, and adjusted 0.782. Generally, it meant that the model explained the dependent variables (PL and SF) quite well.

Table 5

F-Square Overview

	DC	PL	SF
DC		0.213	3.608
PL			
SF		0.476	

Table 5 demonstrates the effect of independent variables on dependent variables. DC has a positive effect on PL with a coefficient of 0.213, a much larger positive effect on SF with a coefficient of 3.608, showing a very strong influence. Meanwhile, PL also affects SF with a coefficient of 0.476, suggesting a moderate influence.

Table 6

Path Coefficients-Matrix

	DC	PL	SF
DC		0.381	0.885
PL			
SF		0.569	

Table 6 illustrates the path coefficients of variables. In the model, the path coefficients were 0.381 (DC→PL), 0.885 (DC→SF), and 0.569 (PL→SF). These coefficients reflected the influencing direction of the relationships in the model.

Analyzing the relationship between variables was to clarify the effect of DC, SF, and PL. The analysis results from the questionnaire are found as follows.

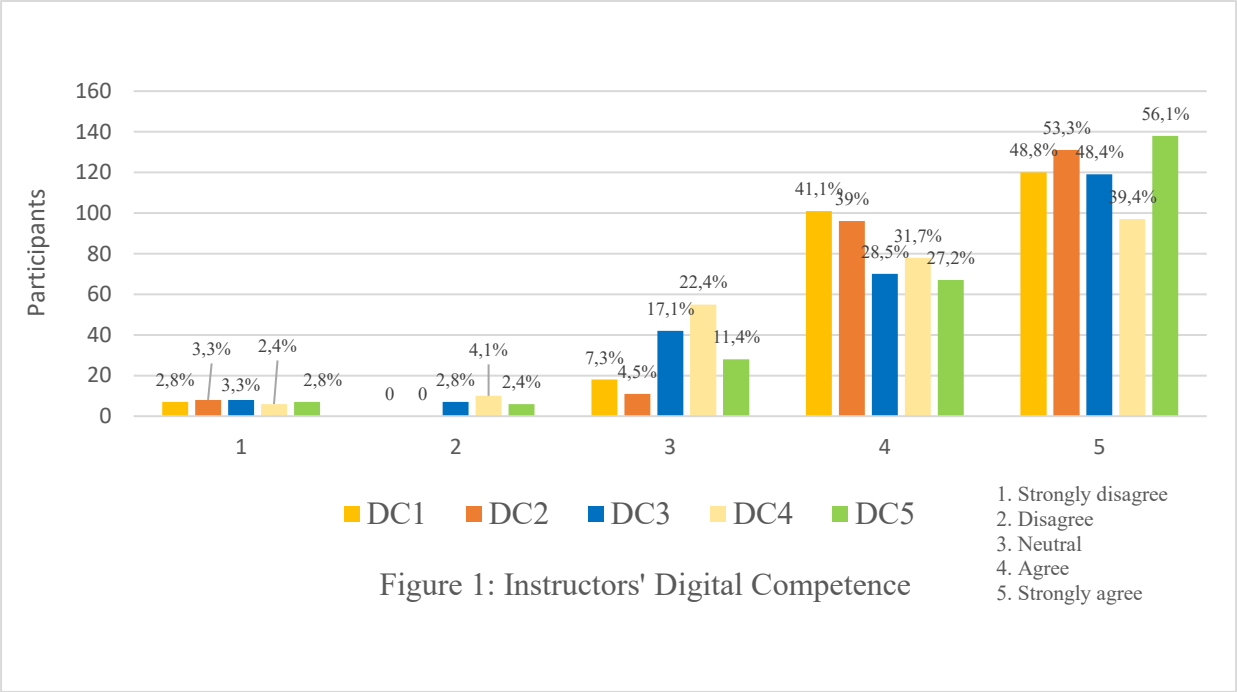


Figure 1: Instructors' Digital Competence

Figure 1 presents instructors' digital competence with five indicators (DC1-DC5). "agreement" and "strong agreement" occupied the highest rate. This proved that instructors were confident about using technology. Particularly, DC2 (selecting digital sources) was the highest, 92,3%; DC1 (effective use of digital tools in teaching) was 89,8%; and DC5 (instructors' support through digital platforms) was (83,3%). While the "Neutral" rate fluctuated from 3,3% (DC1), 2,8% (DC2), 17,1% (DC3 - effective digital classroom organization and management), 28,5% (DC4 - using technology to assess learning outcomes), and 48,4% (DC5). This reflected that some instructors hesitated to use technology. The rate of "disagree" and "strongly disagree" was the lowest in most of the variables, such as DC1 (2,8%), DC2 (3,3%), DC3 (6,1%), DC4 (6,5%), and DC5 (5,2%). Overall, the results proved that instructors' digital competence was highly appreciated in favor of positive agreement, reflecting that instructors were increasingly adapting well to the digital teaching environment and could apply technology in online teaching activities.

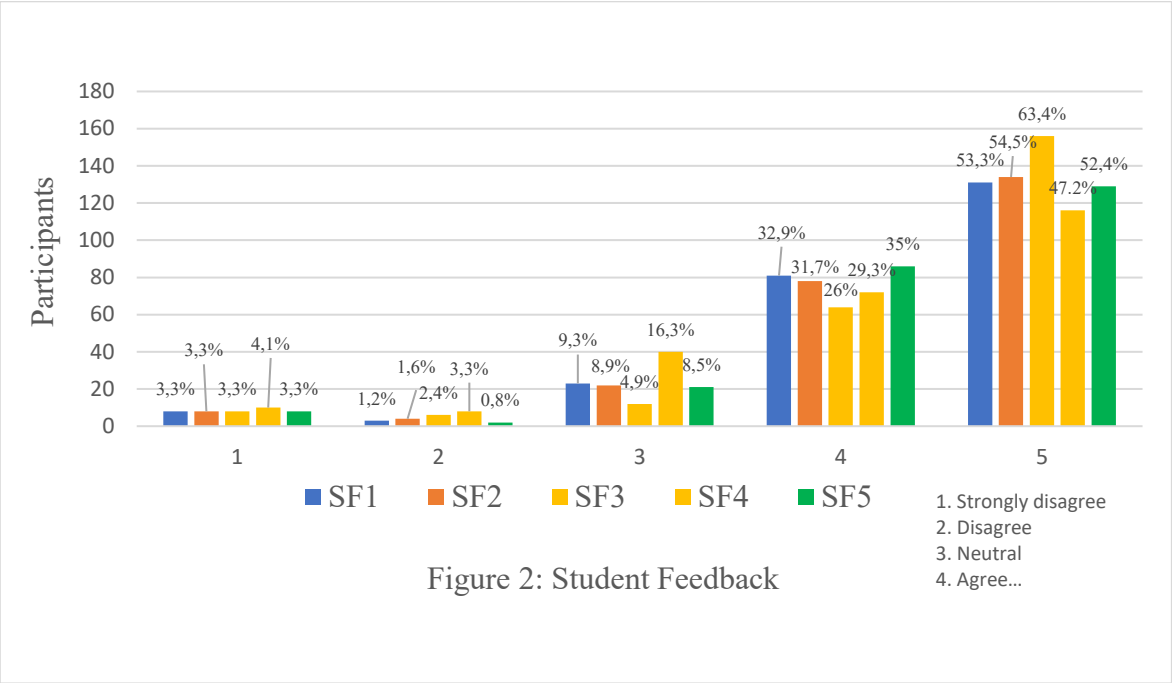


Figure 2: Student Feedback

Figure 2 indicates student feedback in online classrooms through five observed variables (SF1-SF5). Most of participants had positive assessment of feedback at the highest rate of "agree" and "strongly agree", namely 89,4% (SF3 – instructive feedback and learning outcome improvement), 87,4% (SF5 – interaction based on received feedback), 86,2% (SF1 – in-time feedback from instructors), 86,2% (SF2 – feedback helping understand weaknesses and strengths), and 76,5% (SF4 – personalized feedback based on need). In contrast, the rate of "disagree" and "strongly disagree" was very low; SF1 (4,5%), SF2 (4,9%), SF3 (5,7%), SF4 (7,4%), and SF5 (4,1%). Also, the "Neutral" rate was low for variables, ranging from 4,9 to 16,3%. Generally, the results explained that students had positive experience with learning feedback, appreciating the support of the online learning process.

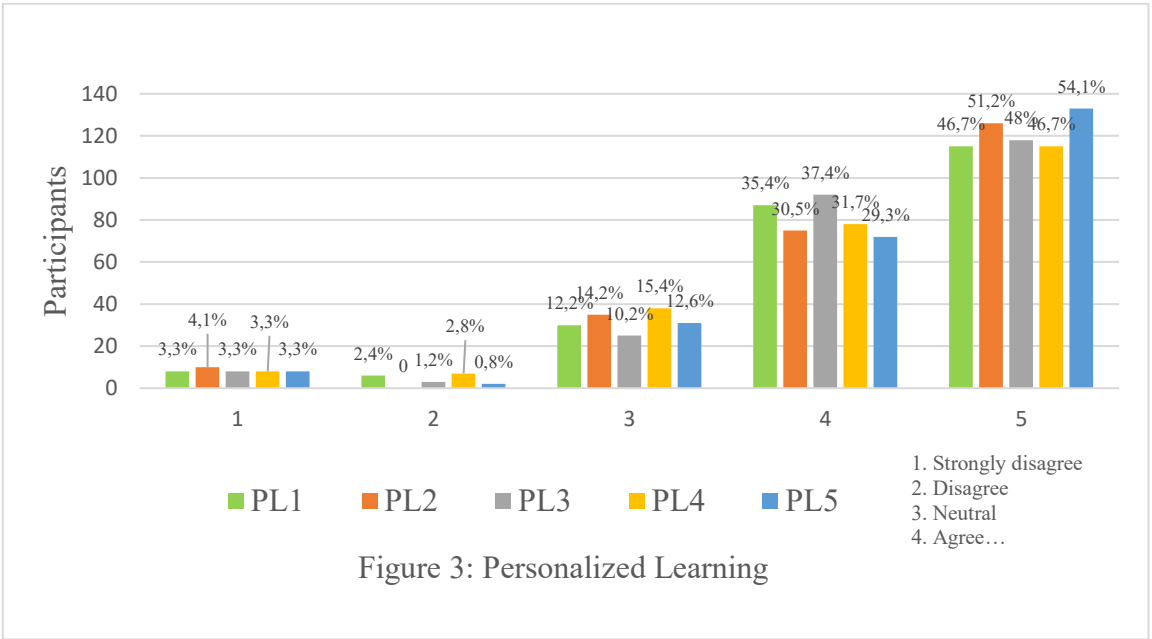


Figure 3: Personalized Learning

Figure 3 illustrates the Personalized Learning (PL1-PL5) framework. According to the results, “agree” and “strongly agree” outnumbered in comparison with other variables, such as 85,4% (PL3 – appropriate materials on the LMS), 83,4% (PL5 – approaching content), 82,1% (PL1 – content appropriateness with individual needs), 81,7% (PL2 – individual learning pace), and 78,4% (PL4 – learning support appropriate with individual competence). These results showed a positive trend in personalized learning, being well-fit with students’ needs and meeting flexibility and individual support in the learning process. The “Neutral” rate fluctuated from 10,2% to 15,4%, reflecting that fewer participants had a clear decision in doing the survey. The very low rates for “disagree” and “strongly disagree” occurred for most of the variables under 5%. There were fewer negative ideas for PL. In overview, the results demonstrated the role of personalized learning in enhancing the online learning experience.

5. Conclusion

In conclusion, there is a positive effect of digital competence on feedback and personalized learning. Instructors’ digital competence consists of selecting digital sources, using digital tools in teaching, providing technical support through digital platforms, organizing and managing effective digital classrooms, and assessing learning outcomes through technology use. Instructors tend to adapt well to the digital teaching environment and apply technology in online teaching activities. The influencing factors are found through the following items. First, HIU instructors’ digital competence positively affects feedback. When instructors have high digital competence, they can provide more rapid and diverse feedback. The forms of feedback include instructive feedback and learning outcome improvement, interaction based on received feedback, in-time feedback from HIU instructors, feedback helping understand weaknesses and strengths, and personalized feedback based on need. Second, digital competence directly affects personalized learning thanks to technology support in designing content and a learning path for each student, selecting appropriate materials on the LMS, approaching content, ensuring content appropriateness with individual needs, maintaining individual learning pace, and providing learning support appropriate to individual competence. Third, HIU students’ feedback has a powerful effect on personalized learning because the quality of feedback helps them adjust the learning process based on individual needs. Finally, HIU student feedback plays a mediating role in the relationship between digital competence and personalized learning. From the above results, the recommendations of the study are that HIU in particular and universities in general should build digital competence training programs and professional development for instructors.

REFERENCES

Betancur-Chicué, V., & García-Valcárcel Muñoz-Repiso, A. (2023). Microlearning for the development of teachers' digital competence related to feedback and decision making. *Education Sciences, 13*(7), 722. https://doi.org/10.3390/educsci13070722

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approach* (5th ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/research-design/book255675

Dang, T. D., Phan, T. T., Vu, T. N. Q., La, T. D., & Pham, V. K. (2024). Digital competence of lecturers and its impact on student learning value in higher education. *Heliyon, 10*(17), e37318. https://doi.org/10.1016/j.heliyon.2024.e37318

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/a-primer-on-partial-least-squares-structural-equation-modeling-pls-sem/book277161

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E., & Mahoney, P. (2019). Conditions that enable effective feedback. *Higher Education Research & Development, 38*(7), 1401–1416. https://doi.org/10.1080/07294360.2019.1657807

Hizam, S. M., Akter, H., Sentosa, I., & Ahmed, W. (2021). Digital competency of educators in the virtual learning environment: A structural equation modeling analysis. *IOP Conference Series: Earth and Environmental Science, 704*(1), 012023. https://doi.org/10.1088/1755-1315/704/1/012023

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign. https://curriculumredesign.org/wp-content/uploads/AI-in-Education-Promises-and-Implications-for-Teaching-and-Learning.pdf

Ilomäki, L., Paavola, S., Lakkala, M., & Kantosalo, A. (2016). Digital competence – An emergent boundary concept for policy and educational research. *Education and Information Technologies, 21*(3), 655–679. https://doi.org/10.1007/s10639-014-9346-4

Instefjord, E. J., & Munthe, E. (2017). Educating digitally competent teachers: A study of integration of professional digital competence in teacher education. *Teaching and Teacher Education, 67*, 37–45. https://doi.org/10.1016/j.tate.2017.05.016

Luketić, D., & Diković, M. (2026). Individualized teaching and personalized learning in higher education: Insights and future directions from systematic mapping revie*w. Trends in Higher Education, 5*(2), Article 45. https://doi.org/10.3390/higheredu5020045

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Pane, J. F., Steiner, E. D., Baird, M. D., Hamilton, L. S., & Pane, J. D. (2017). *Informing progress: Insights on personalized learning implementation and effects*. RAND Corporation. https://doi.org/10.7249/RR2042

Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu. *European Commission*. https://doi.org/10.2760/159770

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Tondeur, J., Van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review. *Educational Technology Research and Development, 65*, 555–575. https://doi.org/10.1007/s11423-016-9481-2

Walkington, C. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology, 105*(4), 932–945. https://doi.org/10.1037/a0031882

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

#81 THE TEACHER AS A FACILITATOR IN CULTIVATING ENGLISH-RICH HOME ENVIRONMENTS: A PARENTAL COACHING APPROACH FOR PRIMARY EDUCATION

Tham Truong Thi [1], Thanh Phan Thi [2]

Abstract

This paper introduces an action research that helped primary students learn English at home with their parents' involvement. In this research, the teacher played a role as a facilitator in coaching the parents to become their children's partners in learning English at home. This bridged the gap between the classroom and home environment for primary students in the context of teaching and learning English as a second language. The research was conducted to investigate how the teacher's guidance to parents and their support would affect students' development in learning English. There were 21 students and their parents in her class participating in the research for four weeks. The teacher used a digital platform (Zalo group) to give guidance and collect data. The teacher used three data collection methods: classroom observation, portfolio and semi-structured interview to ensure objectivity. The preliminary results of the experiment revealed that when the teacher was involved in guiding parents, the home environment became a supportive second-language environment for the child, and this was beneficial for the child's confidence level when they used the target language. It was shown that the teacher-facilitated parental coaching model is an applicable solution for the long-term development of English in the Vietnamese primary school context.

***Keywords*:** English-rich home, teacher as a facilitator, coaching, supporting, primary learners' English development

1. Introduction

The transformation of teaching and learning English from a foreign language to a second language in the Vietnamese education system is considered an innovative step of the globalized human resource development policy in the period of 2025-20245. In the era of rapidly growing digital technology, human resources with English or any foreign language are not prefered anymore. The ability to speak a foreign language is a compulsory criterion in the global labour market.

In the resolution 71-NQ/TW on August 22nd, 2025 about educational and training development, the Politburo focuses on enhancing teachers' and students' English level, step by step, to turn out English as a second language. It leads to an approval through a decision 2371/QĐ-TTg, which was approved "Teaching English as a second language in school from 2025 to 2035, a vision to 2045". To achieve the targets, besides the changes in curriculum programs and the improvement in training teachers at school, English should not be taught only at school, but everywhere in life.

The Research Gap

At present, most students in public primary schools in Ho Chi Minh City are taught English in 3-4 periods per week. This time is not enough for them to practice and immerse to

[1] *Tham Truong Thi, MA., Faculty of Foreign Languages, University of Labour and Social Affairs, Ho Chi Minh City, Vietnam; Email: thamtt@ldxh.edu.vn*
[2] *MA., Faculty of Foreign Languages, University of Labour and Social Affairs, Ho Chi Minh City, Vietnam*

English in addition to classes with a large number of students. After school, students often take part in extra English classes. These classes usually last four hours per week. When they come home, they do not have an environment to practice English because their parents think teaching and learning English is the teacher's responsibility. This causes a language vacuum after school. According to Spolsky (2004), on the theory of Language Policy, one of the domains that should be extremely important for learning language would be family; yet, many Vietnamese families appear to be inactive, which could be attributed to the phenomenon described by Hoover-Dempsey & Sandler (1997) as the "lack of parental self-efficacy". Moreover, many parents are afraid of supporting their child since they are not good at English pronunciation. However, according to Epstein (2011), on school-family partnerships, it is suggested that reinforcing learning at home is significantly pivotal for students to gain success.

To bridge this gap, the coaching role of the teacher should be promoted. Parents are those who manage their children's home time, yet teachers are those who provide the pedagogical methods to help them teach their children.

Research Aim

This action research is conducted with the aim of investigating the effectiveness of teachers' parental coaching. By employing a digital communication platform via Zalo group, the teacher can deliver guidance to instruct parents to create a home-rich English environment, as a natural part of daily life rather than a stressful academic task (Krashen, 1985).

Research Questions

1. How can a teacher coach parents to create an English-rich environment at home for primary students?
2. To what extent does the teacher-facilitated coaching model influence parents' self-efficacy and level of engagement in supporting their children?
3. How does parental support affect students' level of confidence and lesson retention?

2. Literature Review

2.1. Second language acquisition and Language environment

According to Krashen's input hypothesis (1985), human beings acquire a language through understanding comprehensible input. It brings special importance to children who learn a language better through language immersion. Moreover, Lightbown and Spada (2013) emphasize that the length of time of language exposure plays a decisive factor to young learners' success.

Besides that, Sénéchal and LeFevre (2002) developed the home literacy model to prove the approach of unofficial language activities such as listening to music and reading stories at home significantly supports young learners' language acquisition ability at school. An outstanding demonstration from Hart and Risley (1995) shows that a child's language ability doesn't depend on their innate ability, but rather is based on their language exposure in their early family environment. In the second language acquisition context, it is insisted that the family environment is an important extra-class after school.

2.2. Parental participation and Social interaction theory

The zone of proximal development (ZPD) (Vygotsky, 1978) is an important concept in socio-cultural theory. According to the theory, he states that social interaction plays an indispensable part in children' s learning process, influenced by diverse cultures. Having the

same point of view as Piaget, Vygotsky believes children are the most active learners of all ages. ZPD is defined as the zone of potential development, referring to the gap between the realistic development level identified by independent problem solving and the potential development level identified by solving problems when they are instructed by the older or associated with better partners. In this case, parents are more knowledgeable, providing essential scaffolding for children to practice English at home.

In the six types of school-family cooperation, Epstein (2011) identifies that learning at home is a key element. In contrast, parents only participate in their children's learning process when they have sufficient self-efficacy. Along with Epstein (2011), Bandura (1997) shares the same opinion. An action can be taken only when the parents anticipate positive results. Vietnamese parents tend to underestimate their English proficiency.

Based on the theory of scaffolding (Wood, Bruner and Ross, 1976), parents are not required to teach grammar or be another teacher; they are asked to be a scaffolding to step by step elicit and motivate children to study. Therefore, teachers must coach parents to encourage their confidence.

2.3. Teacher as a facilitator

In modern times, the transition from a teacher to a parental instructor means simplifying pedagogical skills so that parents can easily guide their children. Teachers are not the only source of knowledge any more. According to the facilitation theory (Rogers, 1969), a facilitator must believe in others' capacity to learn and to do well. In fact, the role of the teacher is still crucial while parental coaching is a step of the scaffolding model. However, to make it successful, the teacher must ensure to create four factors: confidence, competence, connection and cognition. Traditionally, parents play a role of involvement; nevertheless, this research moves it to empower parents' self-efficacy.

2.4. Digital technology – The bridge connecting teachers and parents

In the context of rapid digital technology development, the relationship between the school and parents needs to be digitalized continually. Parents are not only "supervisors", but also digital partners, corresponding with the teachers in managing, orienting, supporting and evaluating their children's learning process through digital platforms. As Goodall (2016) highlights, technology-mediated engagement successfully transforms parents from external observers into active co-educators. Furthermore, empirical findings by Kuusimäki et al. (2019) confirm that the appropriate implementation of digital communication channels fosters sustained home-school dialogue, effectively integrating parental voices into institutional practices and expanding collaborative learning opportunities.

3. Methodology

3.1 Research design

This action research adopts a qualitative action research design, following the process of planning, acting, observing and reflecting (Hinchey, 2008). The purpose of this research is to improve and enrich the English home environment by coordinating and coaching parents.

The research context is conducted in the author's English classes from grade 1 to grade 5, involving 21 students.

The research objectives and procedures were clearly introduced to all of the participants. By employing a qualitative triangulated framework, namely, observation, Journal analysis and semi-structured interview, I analyzed the authentic data to explore the impact of

teacher-led parental coaching on young students' English acquisition. Additionally, concrete implications were given to help teachers of English improve the quality of their home-school partnerships.

3.2 Participants

There are three groups of participants involved in this action research:

- One teacher (the author) played the role of facilitator and researcher, designing micro-scaffolding exercises, giving directed exercises and observing the process.
- 21 parents acted as home-coaches and followed the teacher's guidance to remind, encourage and interact with their child to help do English exercises at home.
- 21 students from grades 1 to 5 take part in the learning process both in class and at home.

3.3 Data collection

According to Dahlberg & McCaig (2010), practitioner-researchers conduct qualitative research in order to seek in-depth information related to the issue. To raise the reliability and validity of the data, the data was collected around a qualitative triangulation framework utilizing three distinct instruments: observation notes, a portfolio, and semi-structured interviews.

Observation was the first method used in the research to capture real-time classroom behaviours. Dahlberg & McCaig (2010) state that open-type observation is often employed within school settings of action research to analyze children's actions and behaviors. And Finch (2005) further discusses in his paper that observation allows teachers to reflect on the classroom performance. Consequently, observation notes were maintained.

The second was a digital portfolio. The researcher collected and documented qualitative artifacts, including screenshots of parents' text responses, photos and feedback on Zalo groups after delivering the instructions.

The final instrument was semi-structured interviews conducted with 7 selected parents at the end of the research process. Semi-structured interviews are useful for researchers to elicit interviewees' perceptions (Dahlberg & McCaig, 2010).

3.4 Research data and data analyses

Data analysis is a critical phase that transforms raw qualitative evidence into meaningful academic insights. In this research, the data includes transcripts of interviews, notes from classroom observation and a digital portfolio. According to Dahlberg and McCaig (2010), qualitative data analysis follows a fundamental sequence from collecting and coding to categorizing, interpreting, and reporting.

To analyze the data with academic rigor, this study also adopted the systematic thematic analysis framework established by Braun and Clarke (2006), utilizing its phases to identify themes. Simultaneously, the practical, hands-on coding techniques recommended by Hinchey (2008) for practitioner-researchers were integrated into the process.

The integration of these three methodological frameworks was executed through the following structured steps:

Data familiarization and transcription (Dahlberg & McCaig, 2010; Braun & Clarke, 2006): The tape-recorded data from the semi-structured parent interviews were transcribed

verbatim. The researcher repeatedly read the transcripts, classroom observation notes, and Zalo portfolio entries to gain an immersive, in-depth understanding of the text.

Initial code generation and color-coding (Hinchey, 2008; Braun & Clarke, 2006): Raw data segments were assigned initial codes to capture meaningful semantic features. To visually manage and distinguish different kinds of information at this stage, the study employed various practical techniques suggested by Hinchey (2008), such as using distinct colors, symbols, and highlighters. For instance, statements reflecting parental language anxiety were highlighted in one color, while instances of student lesson retention during observations were marked in another. This provided quick and easy visual access to the codes.

Searching and categorizing for themes (Dahlberg & McCaig, 2010; Braun & Clarke, 2006): In line with Dahlberg and McCaig's (2010) phase of categorization and Braun and Clarke's (2006) theme generation, the codes were collated into broader conceptual patterns.

Defining and naming core themes: Three major academic themes directly address the research questions:

- Theme 1: Teacher-led coaching strategies and parental conceptual understanding
- Theme 2: Psychological and behavioral shifts in parental self-efficacy and attitudes across age groups
- Theme 3: Student cognitive outcomes and systemic obstacles to model sustainability

Interpreting and reporting (Dahlberg & McCaig, 2010; Braun & Clarke, 2006): Following the final steps outlined by Dahlberg and McCaig (2010), the analysis was reported by combining practical findings supported by direct quotes from participants to maintain the qualitative trustworthiness of the action research.

4. Results

4.1 Teacher-led coaching strategies and parental conceptual understanding

The first theme of the implementation of teacher-led coaching strategies and how they deliver the messages and show insight into the receiver's perceptions. The data from the Zalo digital portfolio and interview transcripts demonstrated a significant shift in parental pedagogical awareness.

Before the research, the parents just reminded their children to complete their homework without any instructions. Since the teacher introduced explicit and micro-scaffolding instructions to them, the instructions have changed their approach. As shared in an interview:

" *I feel more confident and sometimes I play a game with my child when the teacher instructs the parents."*

This behavioural improvement proves that parental engagement needs clear instructions from teachers. This finding aligns with Rogers (1969) that the educators' role should change from a transmitter of knowledge to a facilitator creating a supportive learning system.

Moreover, portfolio data from the Zalo digital platform revealed a strong correlation between instructional delivery and parent responsiveness: parents who actively supported their children consistently interacted with, "liked," or replied to the teacher's coaching

messages, whereas non-engaged parents rarely reacted. This pattern supports the theoretical premise that technology-mediated communication optimizes home-school partnerships when digital inputs are concise, actionable, and conceptually accessible to non-educators (Kraft & Rogers, 2015; York et al., 2019).

4.2 Psychological and behavioral shifts in parental self-efficacy and attitudes across age groups

According to observation notes and interviews, the coaching model shows that the role of parents changes based on the students' level.

A group of grades 1 and 2: This group had the largest number of parents who often supported them with homework. At this age, students still need clear and detailed instructions from teachers or parents because they lack autonomous learning strategies.

A group of grades 3 and 4: In this intermediate group, most parents move to the role of reminding their children to do homework. Parents provided targeted guidance whenever their children encountered questions.

A group of grade 5: At this senior primary stage, children have formed study habits of independence, even whether their parents remind them to do homework or not. However, by observing them, the researcher realized that the students who finished their coached home assignments were more confident during lessons than those who didn't.

The adjusted parental support according to age proves the scaffolding concept of Wood, Bruner and Poss (1976) and is rooted in Vygotsky's (1978) Zone of Proximal Development. As the learner's competence grows, the parental coach gradually moves from external support to encourage the learner's autonomy.

This finding also supports the perception of Lightbown and Spada (2013) regarding age-related factors in second language acquisition (SLA). Younger learners need more consistent support than older ones in primary school.

4.3 Student cognitive outcomes and systemic obstacles to model sustainability

4.3.1 Impact on Student Lesson Retention and Confidence

Based on the data collection, it can be seen clearly that there are strong discrepancies among students with parental support, those with less support and a group with hardly any support.

The group with entire parental support: In this group, students were reminded to do exercises, directed to review and check after finishing their homework by their parents. The dominant students are Cherry, Xuka, Sunny, and Quyen, who show exceptional cognitive retention of previous lesson content and high confidence in responding to the teacher's questions. Similarly, other students, including Mimi, Hanna and Win, went to class with confidence and superior vocabulary recall ability.

The group with less parental support: In this group, parents just reminded the students generally to do their homework, but gave no scaffolding instructions. As a result, these students (Khoa, Quoc Quan, Minh Quan and An Nhien) just occasionally did homework and revised lessons, leading to inconsistent classroom participation.

The group with no parental support: In this group, the students often felt worried and unconfident, particularly during the initial moments of entering the classroom or when required to check their previous lessons.

These discrepancies evidence the emotional filtering hypothesis (Krashen, 1982). Guided parental support at home creates a psychologically safe environment, minimizing children's anxiety and fear and optimizing language acquisition. These differences also reflect the findings of Hart and Risley (1995) on the profound impact of the early home language environment on learning pathways.

4.3.2 Systemic and Situational Barriers

Through evaluating the results of note-taking and interviews, the paper has identified some main reasons affecting the effectiveness of the model: workplace exhaustion, school workload and parental linguistic anxiety.

The obstacles from work: " I am so tired after work, I just want to let my hair down so I forget to remind my child", said Truc Anh's mom and " I often finish my work late", said mother Thao and Rio's mom. Both of them really want to learn side by side with their children; however, they were attracted by their heavy workload.

"My children's school homework is too much, and they don't have enough time to do their homework in your English extra class", another said.

"I don't know English, so I just remind him to finish his homework", said Phuc's mom.

These empirical findings align strongly with literature on primary EFL education and family dynamics in Vietnam (Hamid & Nguyen, 2016), which underscores that demanding occupational pressures coupled with intensive primary school workloads construct formidable structural barriers for busy parents. Consequently, even though the parental coach model brings positive effects on learners' cognition and emotion when learning English as a second language, it depends significantly on teachers' ability to provide flexible, time-saving and linguistically understandable micro-tasks that maximize parental support.

5. Discussion

5.1 Scaffolding and parents' psychological barriers

The empirical findings indicate that the effectiveness of the model varies depending on the level of parental involvement: active instructional scaffolding, passive behavioural monitoring and complete non-engagement.

According to Wood and associates (1976), accompanying their children in learning at an early age is strongly recommended for parents. When there is only a generic reminder, many students seem distracted. They might struggle with task ambiguity, or sometimes they neglect assignments due to the absence of parental verification.

About 50 percent of parents were not involved in their children's learning process or with a generic reminder. The reasons may not come from a lack of care, but they felt unconfident, low self-efficacy and under pressure of wrong pronunciation or encountered the barrier of pedagogical methods (Bandura, 1997).

The case of student Cherry, who was immersed in full parental support, is a clear demonstration of implementing the right process of distributed scaffolding. This association helps maximize the zone of proximal development (ZPD) according to Vygotsky (1978). Hence, students can actively practice English at home.

5.2 Students' performance and psychology

Observational data indicate that home-school mediation significantly shapes students' psychological readiness in class.

The students who had well-prepared homework came to class with high confidence and active engagement. According to Krashen (1982), when the anxiety diminishes, the affective filter is low. Therefore, children are ready to acquire a second language most successfully and productively.

In contrast, the failure to complete home assignments before going to class creates anxiety for children. It leads to a high affective filter which restricts their language acquisition with timidity and inactivity.

5.3 Learner autonomy

Based on the researcher's observation, the data reflect a linear progression of primary students' autonomy. Most students in grades 1-2 need clear instructions from parents; those in grades 3-4 can do homework themselves and ask their parents if necessary, while those in grade 5 are likely independent. This progress requires flexibility of teachers as a facilitator. In early grades, teachers should coach parents with clear and detailed instructions. Whistle in older grades, teachers should focus on designing learning materials with a student-centered approach, now that they are aware of their responsibilities.

5.4 Practical obstacles

Qualitative interview data reflect the fact that discrepancies in parental engagement are primarily driven by systematic and situational factors.

Because of the dual burden of heavy daily workloads from work and home chores in our modern society, parents sometimes forget to remind their children to do homework. Moreover, besides English homework from the extra English classes, children have to complete the heavy workload of school assignments. Thus, they are not able to manage their time to complete all of them. Additionally, the time balance between learning and relaxing is considered more important to some parents than their high demands on children's learning.

The empirical finding indicates that conventional and time-consuming pedagogical directions may fail in some cases of busy parents. Parents can not change their work schedule; hence, teachers' responsibility is to adapt and streamline instructional delivery methods.

5.5 Recommendations

Roles of Teachers

To balance groups of parents, it is recommended to create short videos that parents can play for children to follow, instead of only written instructions. Making videos can take more effort from teachers but it is worth dedicating.

Teachers can also install automatic reminder functions on Zalo groups with positive encouragement to solve the problem of forgetfulness. Rather than reminding parents in groups, teachers send messages to each busy parent.

Roles of parents

After being instructed by teachers, parents can build an enriched English environment. Parents must be the ones to set up and manage children's learning time, following 3-step scaffolding: Remind (remind children it is time to do homework)-Activate (play videos or instruct if necessary)-Check (check children's answer and reward).

In response to excessive homework, parents often utilize time-budgeting assistance as a crucial scaffolding tool to alleviate psychological distress and minimize student burnout.

Roles of students

Instead of waiting for their parents' instructions, children must learn how to study autonomously to finish their work. Teachers and parents allow them to raise questions in case they address the struggles.

Students are directed to inform teachers initially and ask for help if they forget to do homework to reduce their anxiety and release stress right after that. This will lower their affective filter and keep confidence.

6. Conclusion

This action research was conducted to investigate the effectiveness of parental coaching by teachers as a facilitator. The practitioner-researcher aims to help primary students have chances to practice English in a rich-English home environment. Through data collection methods: observations, portfolio and semi-interview, the research comes to conclusions as follows:

- The research confirms that the act of parental intervening forms students' behaviors and emotions in class. When parents spend enough time supporting students, they go to class with high confidence and acquire the knowledge superior to others in the next lesson. In contrast, those who are not instructed feel worried.

- The model observes a natural progression in children's self-learning abilities. Grades 1-2 rely entirely on external support (currently, 50% of parents provide both reminders and guidance effectively); grades 3-4 begin to work independently and only ask for help when encountering difficulties; while grade 5 students have reached a state of complete self-reliance without needing frequent reminders from parents.

- The breakdown in the support process doesn't stem from a lack of cooperation, but rather from the double pressure of time constraints: parents are busy with overtime work, and students are overwhelmed by the volume of schoolwork.

Understanding parents' barriers and students' anxieties is the key for teachers to create a humane, effective and sustainable educational environment.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dahlberg, L., & McCaig, C. (2010). *Practical research and evaluation: A start-to-finish guide for practitioners*. SAGE Publications.

Epstein, J. L. (2011). *School, family, and community partnerships: Preparing educators and improving schools*. Westview Press.

Finch, J. (2005). Action research: Empowering the teachers. *Pleiades: Journal of Teaching Young Learners of English*, *1*(1), 30–48.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing.

Hinchey, P. H. (2008). *Action research*. Peter Lang Publishing.

Hoover-Dempsey, K. V., & Sandler, H. M. (1997). Why do parents become involved in their children's education? *Review of Educational Research*, *67*(1), 3–42.

Kraft, M. A., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review*, *47*, 49–63. https://doi.org/10.1016/j.econedurev.2015.04.001

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.

Kuusimäki, A.-M., Uusitalo-Malmivaara, L., & Tirri, K. (2019). Parents' and teachers' views on digital communication. *Education and Information Technologies*, *24*(2), 1359–1372.

Nghị quyết số 71-NQTW về đột phá phát triển giáo dục và đào tạo.

Quyết định số 2371/QĐ-TTg phê duyệt Đề án "Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học giai đoạn 2025-2035, tầm nhìn đến năm 2045".

Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford University Press.

Nguyen, T. M. H. (2011). Primary English language education policy in Vietnam: Insights from implementation. *Current Issues in Language Planning*, *12*(2), 225–249. https://doi.org/10.1080/14664208.2011.597048

Rogers, C. R. (1969). *Freedom to learn: A view of what education might become*. Charles E. Merrill.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of literacy: A resolution of a five-year longitudinal study. *Child Development*, *73*(2), 445–460.

Spolsky, B. (2004). *Language policy*. Cambridge University Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

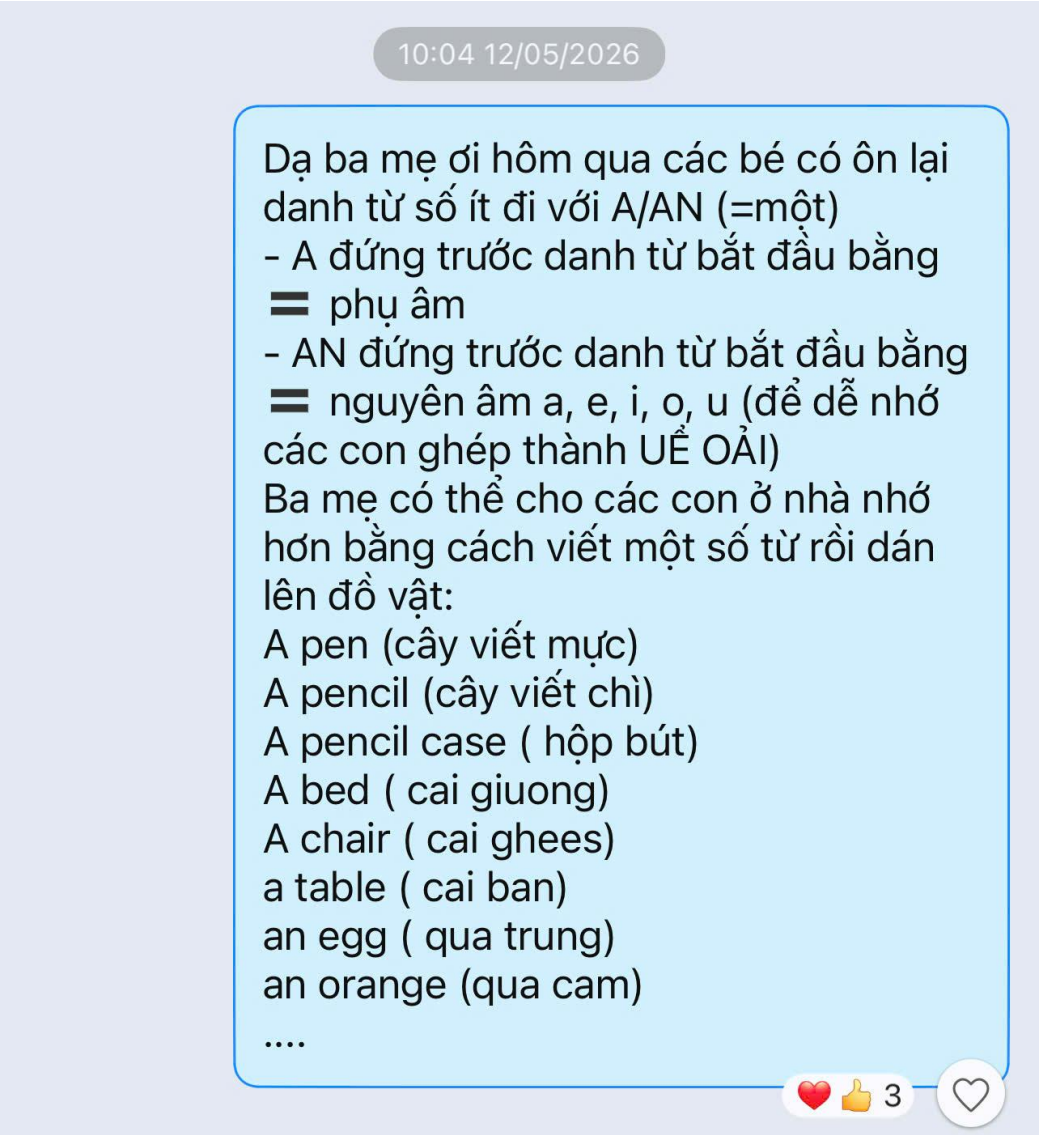
York, B. N., Loeb, S., & Doss, C. (2019). One step at a time: The effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), 537–566. https://doi.org/10.3368/jhr.54.3.0517-8756R

Bionote

Truong Thi Tham and Phan Thi Thanh are English lecturers at the University of Labour and Social Affairs, Campus 2, Ho Chi Minh City. Both of them have been working as lecturers for over ten years. During their teaching time, they have had numerous opportunities to approach innovative teaching and learning methods through the papers of experienced researchers from around the world. They have also written research papers for their university's workshops, in an attempt to find the best way to teach their students.

Appendices

Appendix A: A SAMPLE OF THE TEACHER'S INSTRUCTION



Appendix B : QUESTIONS FOR INTERVIEWING PARENTS

1. How do you usually remind or support your child's homework with instructions from the teacher? (Do you just remind them to do it themselves, or do you sit down and review it with them?)
2. What do you often do when your child doesn't understand a part? (Do you proactively guide and explain it to them, or tell them to ask the teacher the next day?)
3. Can you share any difficulties you often face in reminding your child to do their homework regarding time or work schedules?

4. How much homework do you find in your child's current workload (in the regular curriculum)? Does this workload significantly affect the time your child spends practicing English at home?
5. Are there any barriers in terms of English proficiency or methods that make you feel anxious or hesitant when helping your child?
6. How is your child on days when they complete all their homework before going to class?
7. Conversely, on days they haven't finished homework, do they show any signs of anxiety or share anything with you before going to school?
8. To make supporting your child at home easier, more convenient, and less time-consuming for busy parents, how do you think teachers should change their assignment methods or instruction approaches?
9. Do you have any suggestions for reminder methods (e.g., automated messages, short instructional videos under 1 minute, etc.) to help you avoid forgetting your child's homework schedule?

Appendix C: A SAMPLE OF AN OBSERVATION IN GRADES 1-2

After the teacher assigned homework to the students and gave instructions to the parents, the following happened in the class the next day:

- Cherry, Win, Hana, Mimi: Their mothers reminded them, helped them with their homework and checked their work. These students were very happy and confident when the teacher checked their work, submitting their assignments loudly and clearly. When starting a new lesson, they were also very confident and enthusiastic and learned the new lesson better.

- Khoa: His mother reminded him to do his homework, but he didn't finish it because she didn't give him instructions. He seemed worried when the teacher asked about his homework.

- Quân: His mother went to work and forgot to remind him, and he forgot to do his homework. He said that he also had lots of homework from school. He seemed worried when the teacher asked about his homework.

- Trúc Anh, Nami: Their mothers forgot to remind them to do their homework. These students seemed worried when the teacher asked about their homework.

#82 AI-ASSISTED ESL LEARNING AND STUDENT MOTIVATION: CHALLENGES AND OPPORTUNITIES FOR SOFTWARE ENGINEERING EDUCATION

Do Duc Bich Ngan[42]

Abstract

In Vietnam's software engineering (SE) industry, English proficiency plays an important role in career development and in communicating with international clients. However, for many students in the major, English remains a common obstacle to graduation and future employment. In this context, the rapid growth of Artificial Intelligence (AI) has brought both opportunities and challenges. On the one hand, AI provides valuable support for English learning and academic work. On the other hand, it may reduce students' motivation to learn independently, which leads to possible overreliance on AI tools.

The aim of this paper is to investigate the effects of AI-assisted tools on learning motivation and autonomy of software engineering (SE) students at HUFLIT University. The study mainly employs a quantitative approach, supplemented by a number of open-ended questions to gain deep insights into SE students concerning their learning attitudes and their perceptions of using AI tools for language learning.

While AI tools may develop knowledge, convenience, confidence, and access to learning support, high dependence on AI could reduce students' initiatives and independent thinking in the learning process. Based on the results, the study proposes several AI-assisted learning strategies to support students' motivation, proactiveness, and self-study skills in English learning.

Keywords: AI-assisted language learning (AIALL), computer-based language learning (CALL), English as a second language (ESL), learning motivation, software engineering (SE) education, learner autonomy, interactive AI solutions

1. Introduction

In Vietnam today, English has become a strategic competence that determines career development, as reflected in the Government's Decision No. 2371/QĐ-TTg (2025), which approves the project of "making English a second language in schools for the 2025–2035 period, with a vision to 2045". Particularly, in the field of Science and Technology, students often work with English technical materials and resources. English proficiency enables students to approach up-to-date knowledge. Furthermore, for SE technicians, this not only reduces costs and speeds up project timelines but also helps them stay competitive in the global software market.

At present, there is a mismatch between students' English learning orientations and practical demands. Most software engineering (SE) students tend to focus much on developing

[42] *Information Technology, HUFLIT University, Ho Chi Minh, Vietnam; nganddb@huflit.edu.vn*

technical expertise, whereas underestimating the importance of English learning. As a consequence, some students in this major (SE) are not able to graduate on time or to be well-prepared for their future career. This is not due to a lack of professional competence but because they fail to meet even basic English proficiency (Hong Mai, 2026). In reality, capable Vietnamese software programmers with limited English competence are often restricted to purely technical roles, not being able to take on higher-level responsibilities such as senior developer or manager. Consequently, many IT companies in Vietnam nowadays require their staff not only technical skills but also English proficiency to collaborate directly with international clients for their projects.

Traditionally, English learning has relied mainly on teachers' instructions and later computer-assisted language learning (CALL), which often provides limited personalization or lack of human-like interactions. Recently, AI-assisted language learning tools (AIALL) have marked a significant move, which can personalize students' learning and promote their learning English. At present, several Asian countries, such as China and South Korea have implemented AI-driven platforms with adaptive learning systems and conversational agents, to support English learning through personalized feedback and interaction (Miao et al., 2021, pp. 28–30).

Even though the integration of AI-assisted tools in English learning is constantly increasing, the understanding of how SE students actually use these tools and how such use influences their motivation and autonomy in English learning is still limited. It remains unclear what benefits and challenges they encounter. Currently, in the context of Vietnamese education, while several studies have examined learner autonomy, motivation, and English proficiency among non-English major students (Ho, 2018; Le & Nguyen, 2022; Nguyen & Nguyen, 2020) and some have explored the general benefits of AI tools in English learning (Pham & Pham, 2024; Hoang et al., 2024), limited research has investigated how AI specifically influences the motivation and autonomy of SE students in English learning.

Therefore, this study aims to investigate ESL learning supported by AI-assisted tools among SE students by examining their perceptions, usage experiences, and the effect of AI tools on their learning motivation and autonomy. Below are the research questions of the paper:

1. What are SE students' perceptions of AI-assisted tools in English learning?
2. How do SE students use AI tools for ESL learning, and what benefits and challenges do they experience?
3. How does the integration of AI-assisted language learning tools affect SE students' motivation and autonomy?

Based on the research questions and objectives, two hypotheses are proposed as follows:
Hypothesis 1. SE students' regular interaction with AI-assisted English learning tools is associated with high levels of English learning motivation.

Hypothesis 2. SE students' perceptions of the benefits against the drawbacks of AI tools affect their attitudes toward using AI for learning English.

The study shows an exploratory understanding of AI-assisted ESL learning, focusing on SE students who frequently interact with technology more often. This also helps offer more practical insights and recommendations for teachers and students in integrating AI into English

learning practices. In terms of scope and delimitation, the study only deals with SE students and does not aim to represent students from other majors. It focuses on students' perceptions and experiences rather than objectively measuring English proficiency improvement. In addition, the study examines AI as a language-learning support tool and does not investigate AI applications beyond ESL learning contexts.

2. Literature Review

2.1. Artificial Intelligence (AI) in ESL Learning

AI tools have gradually changed English language learning, shifting from CALL to AIALL. This suggests the new attention to individualized and interactive learning. These new tools are considered to aid in the acquisition of language through personalization, automated feedback, and multi-modal interactions (Selvi et al., 2025). They are currently used in ESL learning contexts, including AI Chatbots (ChatGPT, Google Gemini, …), Automated Writing Evaluation tools (AWEs) (Grammarly, Quillbot, …), and Intelligent Tutoring Systems (ITS). These tools have various applications: Chatbots facilitate active communication and the ability to utilize language in real-world settings (Fryer & Carpenter, 2006; Kohnke et al., 2023), AWEs can also offer instantaneous feedback to learners that leads to better writing accuracy (Stevenson & Phakiti, 2019), and ITS offers personalized instruction that adapts to the abilities of the learner (Holmes et al., 2019). They can individualize content to match the learner's skill level, allow learners to receive instant feedback, and supply various inputs such as text, sound, and visuals, which give learners control over how they learn English, thus providing a more dynamic learning environment. Also, although learning effectiveness is reliant on the extent to which learners critically engage with the output of the tools, studies show that AI tools contribute to an improvement in the quality of learners' writing and encourage sustained practice (Selvi et al., 2025).

Additionally, AI tools function in language learning as writing aids, conversation partners, and personalized tutors which not only promote the learning of English among students but also foster learner autonomy by providing students with the ability to take control over their learning.

As for SE students, English is considered important because they need to access technical documents and interact globally to advance their career opportunities. AI-assisted tools are highly relevant and useful in this context, as they enable learners to engage in interactive practice, receive quick feedback, and develop communication skills beyond traditional classrooms (Huang et al., 2021).

In sum, the tools have gradually proved to be a crucial component of modern ESL learning. However, the extent to which the tools can help learners depends mainly on how learners use them, highlighting the need to examine their impact on motivation and autonomy.

2.2. AI and Learner Motivation and Autonomy

AI tools can influence the learning motivation of students in different ways (Hussain et al., 2025). Obviously, educational environments equipped with AI-assisted tools can bring interactive, personalized, and engaging experiences to learners, so they can enhance learners' motivation. The tools include conversational agents that allow learners to practice communication in low-pressure settings. Moreover, features such as instant feedback and relevant support can strengthen learners' sense of competence. Hence, it is clear that the use of AI tools in language learning may increase

students' willingness to participate and sustain engagement (Fryer & Carpenter, 2006; Huang et al., 2022).

However, AI tools do not increase learners' motivation automatically (Guzmán Alvarado et al., 2025). It depends largely on how they are used and perceived, learners' goals, learning strategies, and level of engagement. If students mostly use AI tools for completing their tasks rather than learning, they may have a limited impact on motivation (Badarudin et al., 2025). Hence, when evaluating AI's role in motivation, it is necessary to consider not only what an AI tool can do but also how learners actually use and respond to it.

According to Liu (2015), motivation and learner autonomy are closely linked in second language learning. It means that motivated learners are more likely to act autonomously, which enhances learners' motivation. Also, Ryan & Deci (2000) point out that learners are more motivated when they can take control of themselves and feel more competent, which fosters intrinsic motivation and deeper engagement. It can be derived that a student who feels autonomous but incompetent may feel unconfident or lack motivation. As a result, students can advance their language proficiency in a way that promotes their autonomy and competence, which in turn strengthens their motivation. Furthermore, AI-assisted learning environments can support the development of autonomy by enabling learners to set goals, monitor progress, and engage in self-directed learning (Yang, 2026). Tools such as chatbots and writing assistants allow learners to practice continuously without constant teacher intervention, encouraging greater responsibility for their learning.

In the Vietnamese context, some empirical studies consistently emphasize the importance of learner autonomy in improving English proficiency, particularly among non-English major students. Ho (2018) found that higher levels of autonomy are associated with better learning outcomes (scores) in English language courses. Similarly, Mohebbi (2024) highlights the need to promote autonomy to help learners adapt to AI-supported learning environments.

In conclusion, AI has the potential to improve both learner motivation and learning autonomy; however, its impact depends on how learners engage with it. This suggests that integrating AI in ESL learning involves not only the use of AI tools but also guidance that encourages learners to study purposefully and independently**.**

2.3. Opportunities and Challenges of AI in ESL Learning

Current studies show that AI-assisted tools offer significant opportunities for English learning by enabling personalized learning that adapts to learners' needs and skill gaps, shifting away from fixed learning paths (Selvi et al., 2025). They also support the enhancement of dynamic and interactive content, which allows learners to deal with materials actively and improve both engagement and learning outcomes as compared with traditional approaches (Holmes et al., 2019). In addition, AI brings about more formative and collaborative feedback, where both AI systems and teachers provide immediate and targeted support to correct learners' misconceptions (Holmes et al., 2019). Furthermore, AI develops student creativity by exposing learners to a wide range of views and informational diversity, which in turn facilitates knowledge acquisition, absorption, and combination processes (Wang et al., 2026).

Despite its potential opportunities, the integration of AI in ESL learning offers several challenges that need careful consideration. The first concern is the accuracy and reliability of AI

content generated. Though AI can produce language with correct grammar, due to a lack of contextual and cultural understanding, it may lead to superficial or misleading interpretations (Creely, 2024). As a result, learners who rely on AI outputs without critical thinking may develop an incomplete understanding of language use.

The second challenge is overreliance on AI tools among learners, which is viewed as “a significant risk” for learners (Zhai et al., 2024). Where learners overly depend on AI to write, translate, or solve problems, they are less likely to engage in critical thinking and language production activities. This is especially true for non-English major students who may not have strong self-regulation skills (Le & Nguyen, 2022). In short, the overuse of AI can negatively affect the level of critical thinking, turning learners into passive recipients of information rather than being actively involved in the learning process (Dergaa et al., 2023; Kasneci et al., 2023; Tian & Zhang, 2025). This implies that overreliance on AI can prevent learners from analyzing and evaluating the available information independently.

Another major ethical concern arising out of the use of AI includes issues related to data privacy, academic integrity, and responsible AI use. Learners are at a risk of revealing sensitive information about themselves while interacting with AI systems. Also, the ease in creating content makes it easy for learners to plagiarize and misuse the AI generated data (Pikhart et al.,2025).

In conclusion, AI is helpful in ESL learning, but its effectiveness depends on how thoughtfully it is used. Rather than replacing the learner’s work, AI must be viewed as an enabling tool that helps to develop critical thinking and independent learning among students.

2.4 Research Gap

An increasing number of studies have investigated English learning among non-English majors in Vietnam. Their subjects revolve around the learner's motivation, autonomy, and language proficiency. Some previous studies have indicated the connection between learners' autonomy and English skills (Ho, 2018). However, most learners have always relied heavily on teacher instructions while struggling to develop their self-regulation skills (Nguyen & Nguyen, 2020; Le & Nguyen, 2022).

Other prevalent difficulties faced by the learners are demotivation, lack of practice opportunities, and lack of confidence in communication (Nguyen, 2024; Nguyen & Pham, 2025). In recent years, there has emerged new research investigating the impact of using AI on learning English. The studies commonly reveal positive results, which include increased language proficiency, motivation, learner autonomy, and opportunities for autonomous learning (Pham & Pham, 2025; Hoang et al., 2025). Experimental data demonstrate the benefits of applying AI-based technologies, such as Chatbots, including the enhancement of speaking skills and increased confidence (Nguyen, 2024).

However, this area remains underexplored when talking about SE students. Bear in mind that they often have to deal with various technologies in their work, and that their studies are associated with specific English demands. Moreover, they have greater advantages in terms of AI literacy compared with students from other majors. With their technical background, many of them may handle the tools beyond ordinary usage by adapting, integrating, customizing and personalizing English learning applications. Thus, the study aims to fill the gap in the existing body of knowledge by investigating AI applications in SE students' English study.

3. Methodology

3.1. Research Design

The study probes into how AI influences SE students' feelings, attitudes, and behaviors. It tries to understand the main concern about whether AI tends to get them motivated or gets in the way of their motivation and learner autonomy, especially in the context where English is not their major.

Data were gathered using online surveys that aimed at identifying variables, including AI usage frequency among SE students, their attitude towards AI, their level of dependence on AI, as well as their perspectives about how AI influences their motivation levels. Likert scale questions were developed, while multiple-choice questions were converted into binary variables. In addition, there are seven open-ended questions to understand the reason why SE students utilize AI tools, what they view as its learning benefits, and what problems they have to deal with. Also, there are some questions focusing on students' views about what needs to be done. Content analysis was employed to identify recurring themes associated with learner autonomy and SE students' expectations from AI-based learning environments.

a. Participants

The sample consisted of 134 undergraduate Software Engineering majors at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). All were in their third or fourth year. The group was viewed as highly representative for some reasons:

- Their major is Software Engineering of the Faculty of Information Technology. Most of them were studying specialized subjects of the third and fourth year and were familiar with English technical materials. However, they were also under pressure to satisfy the university's English graduation requirements while keeping up with demanding projects in their program.
- In terms of technology, all participants had already used AI tools before the study, and they used those tools in different subjects. This may make the information collected from the sample relevant and suitable for the research.
- Gender factor is not what the paper refers to, as the SE students are nearly male-dominated, with female students accounting for under 10% of the participants.

b. Instrumentation & Data Collection

Data for this paper were collected through a structured questionnaire consisting of 25 items for an online survey. The instrument was designed to cover five interrelated dimensions.

- **Group 1**: A list of items (Items 1–7) enquires about students' awareness of learning outcomes and the importance of English. This is to assess students' grasp of the university requirements of learning outcomes and the career expectations they have about the labor market.
- **Group 2**: Difficulties and Barriers (Items 8–10), this group of items is designed to address the challenges and pressure students face when they have to work with English materials in technical subjects, especially common difficulties in English for Special Purposes (ESP) courses.

- **Group 3**: Interaction with AI Tools (Items 11–16), examined why students use AI at university, e.g., to obtain code explanations, translate documents, correct grammar, or receive support for English learning, and it also considered their prompt engineering abilities.
- **Group 4:** The group is about the impact on students' motivation and autonomy (Items 17–19), which is to assess changes in learning attitudes and the extent to which students rely on technology.
- **Group 5:** Students' recommendations (Items 20–25): The students collected qualitative input on AI use across language skills, and they also recorded students' views on how teaching practices might need to change for teachers of specialized subjects and teachers of English under evolving learning conditions like changing learning environments, technologies, AI use, or online learning.

The questionnaires were disseminated in April 2026. The survey was designed to help objectivity in responses, and the student participation was voluntary and anonymous.

3.4. Data Standardization and Analysis

Before analysis, responses collected via Google Forms were cleaned and standardized to ensure the dataset conformed to the needs of statistical modelling. Two complementary procedures were applied.

3.4.1. Processing of Quantitative Data:

All quantitative responses were coded before descriptive and correlational analyses were performed. The categorical variables were then transformed into numerical forms for statistical handling. For example, academic year and proficiency level were assigned numerical codes (e.g., YEAR: 1 = third-year students, 2 = fourth-year students). For analysis purposes, responses to multiple-choice questions were converted into binary variables (0 = not selected, 1 = selected) to calculate the frequency and percentage of each category. Finally, indicators of learner behavior were created by combining related responses using average or total scores, depending on the type of items.

3.4.2. Content analysis of open-ended responses (Qualitative)

The analysis was applied to open-ended responses, which was carried out in three steps.

- The first step was to clean the text data by removing duplicate entries and ambiguous responses.
- The second step was thematic coding, where responses were categorized into core themes regarding students' perspectives on AI tools and pedagogical proposals.
- The third step involved keyword extraction from students' responses for improvement suggestions. A frequency-based scan was performed to identify the most common student requests, which served as the basis for the recommendations in Section 5.

4. Results and Discussion

4.1. Results

4.1.1. Current Perceptions and Challenges (Score Overview)

As shown in Table 1, SE students report a very strong awareness of English as both a graduation requirement and a resource for future employability. At the same time, the growing presence of AI tools is associated with a noticeable change in how students approach English language-related work.

Table 1

Level of awareness of English learning via AI

Variable (Scale)	Mean	Std. Dev.	Interpretation
IMPORTANCE (1-5)	4.37	0.95	Students show a very high awareness of the importance of English.
SUBJECTIVE (1-5)	3.61	1.12	Students tend to believe that deep understanding of terminology is necessary rather than relying solely on AI.
DIFFICULTY_SCORE (0-6)	2.70	1.79	Moderate level of difficulty in accessing technical English.
READING LEVEL (1-3)	1.67	0.66	Most students feel pressure when reading original materials without translation tools.

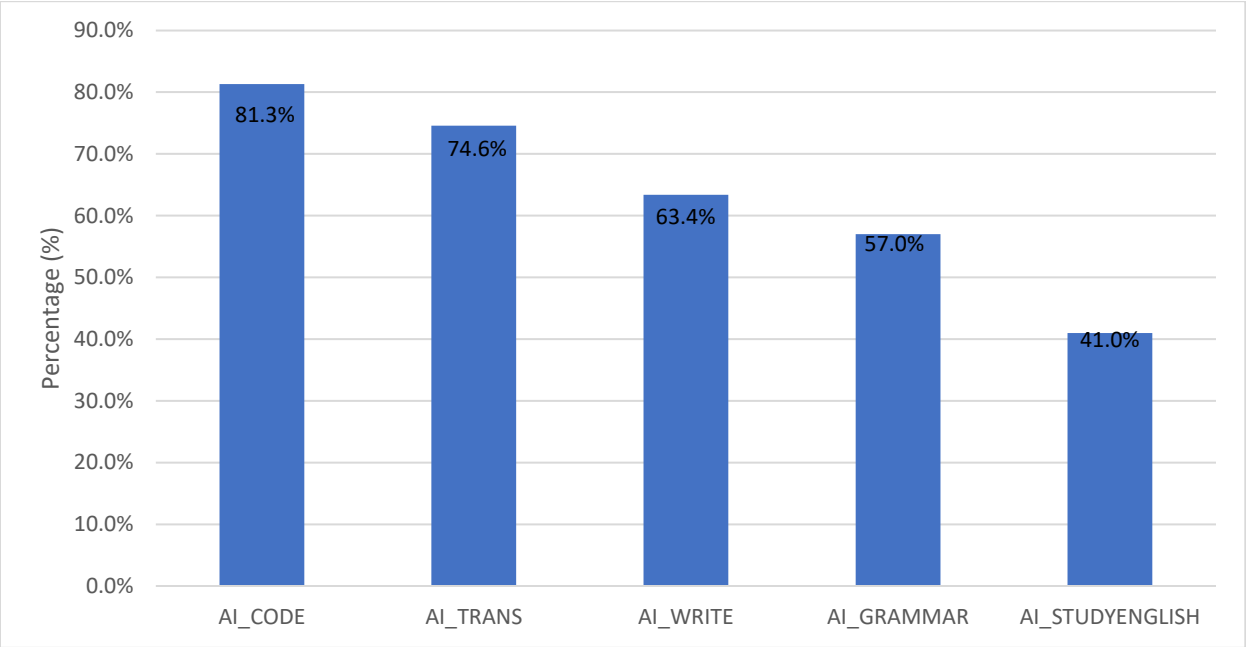
Note: N = 134. Higher mean scores in IMPORTANCE and SUBJECTIVE indicate stronger agreement. DIFFICULTY_SCORE represents the cumulative index of perceived barriers, while READING LEVEL reflects the perceived necessity of translation tools during independent reading.

4.1.2 AI Usage Purposes

The patterns of use indicate that AI serves purposes extending well beyond basic translation. Figure 1 reports that the most frequent application is the explanation of code logic and technical terminology (AI_CODE: 81.3%); English-Vietnamese translation to support reading comprehension (AI_TRANS: 74.6%), assistance with technical report writing (AI_WRITE: 63.4%) and grammar review or practice (AI_GRAMMAR: 57%). By contrast, only 41% of respondents indicate that they use AI explicitly to learn English itself (AI_STUDYENGLISH), for instance, through grammar work or active vocabulary development.

Figure 1

The primary purpose of using AI



4.1.3. Impact of AI on Learning Motivation (Motivation Change)

Table 2

Learning Motivation and Levels of AI Dependency (N = 134)

Component	Category	n	%	Interpretation
Motivation change	Increased motivation	63	47.0	AI reduces barriers and fosters new learning interests.
	No change	44	32.8	AI is perceived as a normal, supportive tool.
	Decreased motivation	27	20.2	Indicates reduced effort due to reliance on AI.
Dependency level	Full autonomy	14	10.4	AI is used only when necessary.
	Companion	76	56.7	AI supports learning while students retain control.
	Dependence (partial)	34	25.4	Students may struggle without AI support.
	Dependence (total)	10	7.5	
Verification behavior	Always/Frequently	82	61.2	Reflects stronger autonomy and critical thinking.
	Sometimes/Never	52	38.8	Risk of accepting inaccurate AI-generated information.

Note. N = total number of respondents.

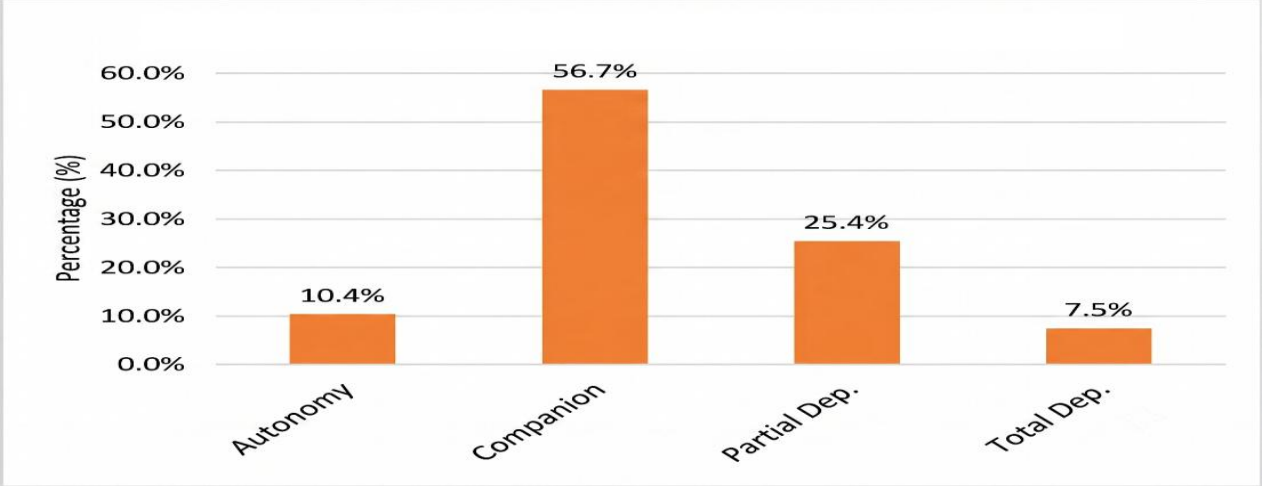
From the empirical evidence presented in Table 2, it can be seen that a variety of attitudes exist concerning motivation, as follows**:**

- Increasing motivation (47%): In this group, the motivation levels are high since AI reduces their anxiety about dealing with complex documents.
- No change (32.8%): The participants see AI as a neutral tool, and therefore, it does not influence motivation levels.
- Decreasing motivation (20.2%): This is an alarming group because the students become accustomed to thinking, "AI is doing everything, and thus I don't have any reason to study."

4.1.4. AI Dependency Levels

Figure 2

Level of Dependency on AI in Language Learning



The distinction between "companionship" and "dependency" is particularly noteworthy:

- 56.7% of participants view AI as a companion
- 25.4% acknowledge partial dependency
- 10.4% show a high level of autonomy
- 7.5% report total dependency

Further analysis shows that dependency and motivation are related. Students who view AI as a "companion" have the highest motivation score (2.51), and those who are "totally dependent" have the lowest (1.70). This suggests that AI improves motivation only if learners have control over its application.

4.1.5. Autonomy vs. Dependence

Table 2 shows that the majority of students (56.7%) view AI as a 'companion', but a significant percentage 32.9% say they are partially or fully dependent. More specifically, the dependence on AI can be seen in Table 3, the students had to encounter high pressure when they read English materials without AI assistance.

The mean pressure score of 1.7 suggests a potential risk of overreliance in ESL learning contexts.

Table 3

The Level of Pressure Experienced by SE Students

State of Pressure	Level (3-point Likert scale)	Description	Frequency (n)	Percentage (%)
Students feel a high level of pressure	1	High Pressure when reading English materials without AI assistance	58	43.28%.
Moderate Pressure	2	Students feel a moderate level of pressure when reading English materials without AI assistance.	62	46.27%
Low Pressure	3	Students feel a low level of pressure when reading English materials without AI assistance.	14	10.45%
Mean Pressure Score	**1.7**			
Total			134	100%

Note. *The scale ranges from 1 (High Pressure) to 3 (Low Pressure). A lower mean score indicates higher perceived pressure*

.

4.1.6. Students' evaluation and proposal of an AI application in English learning

Analyzing the content of open-ended responses resulted in the following themes:

- The first thing is that AI is perceived as a tool that makes complicated information easily comprehensible and removes cognitive barriers. The majority of students indicated that the application of AI makes learning easier. One student wrote, “AI does the work quite well”, another that “AI can translate according to the specific context.” Other responses emphasized its speed, convenience, and utility across all aspects of language acquisition, frequently describing AI as a personal *"tutor."*

- Second, many students raised critical thoughts regarding potential overreliance on AI. Specifically, the necessity of cross-checking of AI outcomes was expressed by several learners. One of the students wrote: “AI might not be 100% accurate for expert languages”, another stated that “students should only keep a part and adjust the others”.

- Third, a tendency to learning in the context where AI is more broadly applied became clear among the learners; there was a clear move toward “practical”, “AI-integrated” and “less theory-oriented” learning. Students suggested project-based learning and more practice-oriented instruction, and the provision of explicit training regarding proper AI utilization, focusing on how to maximize efficiency while avoiding overreliance.

- Lastly, they suggest integrating AI into teaching and learning, improving teachers' AI literacy, and guiding students on using AI effectively. They also stress the need to make the curriculum more practical and relevant to students' majors, to promote learner proactivity, and to adopt more competency-based and project-oriented assessment approaches.

Table 4

Students' feedback and Standardized Themes

Standardized Themes	Percentage	Recommendations
THEME_AI_ROLE	23.13%	Integrate AI into teaching and learning
THEME_TEACHING	29.85%	English teachers should be AI-literate; subject lecturers should guide prompt use, promote innovation in teaching practices.
THEME_CURRICULUM	9.70%	Incorporate practical elements and integrate skills into major-related projects.
THEME_STUDENT	10.45%	Students should be proactive and develop a full range of skills
THEME_ASSESSMENT	17.16%	English assessment should be practical, competency-based, and integrated into major projects.
THEME_GENERAL	9,70%	No specific opinion

5. Discussion

5.1. Students' perceptions of AI-assisted English Learning

SE students have a high perception of the importance of learning English, rating its significance at 4.37. These results are consistent with the relevance of English in accessing technical literature and collaborating globally (Chin & Lah, 2025). At the same time, changes in learning practices are currently happening in the direction of introducing AI to learning (Selvi et al., 2025).

According to open-ended survey responses, students regard AI as a tool making the process of learning less stressful and simplifying access to complex information. They often describe it as "fast" and "convenient," even comparing it to "tutors." Therefore, it appears that AI plays an important role in students' interaction with English in the educational process. Verification of AI outputs was mentioned frequently by many students as necessary, especially for specialization, indicating their critical thinking skills (Benson, 2011; Yang, 2026).

Figure 1 shows that students who find AI useful tend to integrate it into various learning activities, including understanding technical concepts, supporting reading comprehension, writing reports, and reviewing grammar. Moreover, students' frequent verification of AI-generated information indicates an awareness of the need for critical evaluation, reflecting certain aspects of learner autonomy. However, the findings also suggest that perceived usefulness do not always lead to independent learning behaviors. While 56.7% of students consider AI a learning companion, a proportion of students still experience partial or complete dependence on AI.

5.2. AI Use: Performance vs. Learning

As shown in Figure 1, AI is mainly used to perform tasks aimed at code explanation, translation for reading, and writing. Thus, the presented results are consistent with its nature as a supplementary aid in the context of educational institutions, which suggests that students prefer performance to long-term language learning (Tian & Zhang, 2025). This is further confirmed by the answers offered by the respondents, focusing on efficiency and getting tasks done rather than developing language skills more deliberately. Although such an approach is appropriate in the case of students in technical fields (Chin & Lah, 2025), it reflects the gap between its potential as a pedagogical technique and its implementation. Previous studies have also found that such an approach to the matter may limit students' opportunities to develop comprehension (Dergaa et al., 2023; Kasneci et al., 2023).

5.3. AI, Motivation, and Learner Engagement

There is evidence of contradictory results concerning motivation among the respondents. While 47% of them report a rise in motivation, 20.2% of them mention reduced motivation. So, it is claimed that the results are partly aligned with the first hypothesis: the more frequent SE students use the AI tools, the higher motivation they may show. This is mainly because AI decreases anxiety, provides instant help, and boosts their confidence to fulfil English-related assignments. However, even though SE students often use the tools, it does not automatically lead to high motivation, as some SE students admitted they had less effort when AI implemented the learning tasks for them.

Within the framework of Self-Determination Theory, AI may serve as a motivational factor by promoting perceived competence and helping to complete the task (Ryan & Deci, 2000). According to the qualitative results, AI helps students to eliminate fear and intimidation associated with learning, thereby making the process motivating. However, some students report the opposite –specifically, that AI makes it possible to reduce efforts because the tools "execute the work quite well." In other words, motivation seems to be affected by the manner of the technology's use rather than the technology itself (Dörnyei, 2001).

Therefore, the results partially support the hypothesis concerning the relationship between regular interactions with AI-assisted learning tools and high motivation. Table 2 suggests that AI can reduce learning barriers, decrease anxiety when dealing with complex English materials, and encourage students to engage more with English learning. Nevertheless, some students experienced decreased motivation due to overreliance on AI, which may reduce their effort to develop their own skills.

5.4. AI Dependency and the Autonomy–Motivation Relationship

The findings raise issues concerning autonomy and dependency when SE students work with AI. According to the participants' responses, although they are aware that AI outputs may contain inaccuracies (they often verify or correct them), many still rely on AI to complete their tasks. Furthermore, the data suggests a relationship between autonomy and learning motivation: high autonomy corresponds to relatively high motivation ($M = 2.51$), while high dependency is associated with low motivation ($M = 1.70$). In other words, these findings are consistent with Benson's (2011) view regarding autonomy as a central role in sustaining motivation and effective learning. That means SE students with higher autonomy demonstrate higher motivation compared with those who rely heavily on AI.

5.5. Challenges and Opportunities

AI clearly helps reduce language barriers and enables better access to technical materials. Students often describe it as fast and convenient, making it easier to understand and complete tasks. However, moderate difficulty levels (M= 2.70) and low confidence in independent reading (M = 1.67) show that students still rely on AI for support. The qualitative responses also point to the risk of putting in less effort, especially if AI is used instead of thinking critically. The students want more hands-on, AI-integrated instruction that focuses less on theory and provides help with prompt writing and effective AI usage. This indicates that the main challenge lies not in accessing AI but in promoting intentional, critical, and self-directed use in language learning environments.

The findings strongly support the second hypothesis concerning the relationship between SE students' attitudes toward AI and their awareness of both advantages and drawbacks. Many students view AI positively because it makes learning faster, more convenient, and less stressful, describing it as a personal "tutor" that helps simplify complex information. At the same time, students demonstrate awareness of AI limitations, particularly concerns about inaccurate information and overdependence. Several respondents emphasized the need to verify AI-generated content and adjust AI responses rather than accept them completely.

6. Conclusion & Recommendations

6.1. Conclusion

In accordance with the findings, SE learners are aware of the usefulness of English language in academics and professional career, yet they do use AI for certain tasks and not more often for learning purposes. In fact, AI mainly helps them in understanding technical content and doing academic work. Thus, it stresses the practical use of English in the SE field.

In addition, the impacts of AI on motivation among the SE students was not the same. Further evidence suggests that whereas AI may reduce anxiety related to learning, it may promote passivity to create 'a passive receiver'- not an active learner. As a result, it is important to emphasize and help SE students be aware of AI paradox in the way it can prove to be helpful as well as harmful. When students use AI effectively, they will find it easier to be more independent and self-motivated, which may have significant benefits for them. Using it excessively and unconsciously, on the contrary, may prevent learners from engaging fully and developing the ability to learn independently.

In conclusion, the successful integration of AI in English learning relies not just on access but also on how it is utilized. These results indicate the need for guided usage of AI, focusing on encouraging critical thinking, independence, and a learning-oriented approach, along with IT efficiency of SE students.

6.2. Recommendation

From the above results, recommendations are made to improve the learning of English by AI-supported tools for SE students, SE lecturers and lecturers of English.

First, AI-supported tools should be used in the learning process of English. AI should be utilized by addressing both students' academic and professional needs; therefore, students should be asked to use it not merely as a tool for performing certain tasks but as a part of the learning process to promote learning the language, thinking process and engagement.

Second, in terms of English courses designed in the curriculum for SE students, they should be revised to ensure a more appropriate alignment between expected English learning outcomes and the practical demands of SE students in their future career. Greater concentration should be placed on students' communicative competence and their use of English in SE context. Furthermore, the IT Department should provide SE students with practical training in prompt engineering to apply AI-assisted tools for learning English more effectively. SE lecturers could encourage students to initiate AI-based projects, guide them in customizing AI tools for English learning, design and personalize learning software to meet their individual learning needs. Moreover, task design in curriculum should be oriented to integrating AI with English learning.

Third, traditional evaluation is not aligned with AI-supported learning. Since AI tools currently are capable of producing simply memorized or template responses readily, assessment must be adapted to examine how students employ AI tools in an informed, appropriate, and effective manner when completing learning tasks.

Fourth, it is crucial to emphasize that AI is a supportive tool, not a replacement for learning. Thus, SE students need to develop the skills for independent and self-guided learning. Also, to prevent overreliance on AI, it is important to encourage students' active participation in the learning process.

Finally, Vietnamese universities should provide a clear policy on how AI can and should be used ethically and effectively. Such a policy should address academic integrity and offer guidelines to help students use AI responsibly.

Acknowledgements

I would like to thank my partner, Nguyen Van Chuong (MA) and my son, Nguyen Do Nguyen (ECE2023, Vietnamese-German University) for their assistance with English language revising and editing my manuscript.

I would also like to thank 134 software engineering (SE) students at HUFLIT University for participating in the survey for this research.

Last but not least, some AI-based tools including Grammarly and Chat GPT have also helped me a lot with checking grammar, spelling errors, and sentence structures.

REFERENCES

Alvarado, M. V. G., & Andrade, S. S. N. (2025). The impact of artificial intelligence on English language learning: A systematic review of tools, methods, and outcomes in language skills. Runas. Journal of Education and Culture, 6(12), e250287. https://doi.org/10.46652/runas.v6i12.287

Badarudin, B., Parhanuddin, L., Tohri, A., & Suhardi, M. (2024). The effect of AI in education on student motivation: A systematic literature review. Journal for Lesson and Learning Studies, 8(1), 1–10. https://doi.org/10.23887/jlls.v8i1.91141

Benson, P. (2011). Teaching and researching autonomy in language learning (2nd ed.). Routledge.

Chin, M. K. S., & Lah, C. F. C. (2025). The role of AI in supporting English language learning in technical and vocational education. GADING Journal for the Social Sciences, 28(2), 251–262.

Creely, E. (2024a). Exploring the role of generative AI in enhancing language learning: Opportunities and challenges. International Journal of Changes in Education, 1(3). https://doi.org/10.47852/bonviewIJCE42022495

Dergaa I, Chamari K, Zmijewski P, Ben Saad H. (2023). From human writing to artificial intelligence-generated text: Examining the prospects and potential threats of ChatGPT in academic writing. Biology of Sport, 40(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

Dörnyei, Z. (2001). Motivational strategies in the language classroom. Cambridge University Press. https://doi.org/10.1017/CBO9780511667343

Enkelejda Kasneci; Kathrin Sessler; Stefan Küchemann; Maria Bannert; Daryna Dementieva; Frank Fischer; Urs Gasser; Georg Groh; Stephan Günnemann; Eyke Hüllermeier; Stephan Krusche; Gitta Kutyniok; Tilman Michaeli; Claudia Nerdel; Jürgen Pfeffer; Oleksandra Poquet; Michael Sailer; Albrecht Schmidt; Tina Seidel; Matthias Stadler; Jochen Weller; Jochen Kuhn; Gjergji Kasneci. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Fryer, L. K., & Carpenter, R. (2006). Bots as language learning tools. Language Learning & Technology, 10(3), 8–14.

Hoàng, K. T., Lê, T. T. H., Lưu, V. H., & Lê, T. T. H. (2025). ẢNH HƯỞNG CỦA CÁC CÔNG CỤ TRÍ TUỆ NHÂN TẠO LÊN QUÁ TRÌNH TỰ HỌC CỦA SINH VIÊN KHÔNG CHUYÊN NGỮ TRƯỜNG ĐẠI HỌC HỒNG ĐỨC. Tạp chí Khoa học Trường Đại học Hồng Đức – Khoa học Giáo dục và Quản lý, 80(10), 189-199. https://hdujs.edu.vn/index.php/hdujs/article/view/824

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign https://curriculumredesign.org/our-work/artificial-intelligence-in-education/

Hồng Mai. (2026, Ngày 11 tháng 1). Tỷ lệ sinh viên khối kỹ thuật tốt nghiệp đúng hạn còn thấp, trường đại học liệt kê các nguyên nhân. Tạp chí điện tử Giáo dục Việt Nam. https://giaoduc.net.vn/ty-le-sinh-vien-khoi-ky-thuat-tot-nghiep-dung-han-con-thap-truong-dai-hoc-liet-ke-cac-nguyen-nhan-post257142.gd

Hooshyar, D., Šír, G., Yang, Y., Kikas, E., Hämäläinen, R., Kärkkäinen, T., Gašević, D., & Azevedo, R. (2025). Towards responsible AI for education: Hybrid human-AI to confront the Elephant in the room. Computers and Education: Artificial Intelligence, 9, Article 100524. https://doi.org/10.1016/j.caeai.2025.100524

Huang, W., Hew, K. F., & Fryer, L. K. (2021). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. Journal of Computer Assisted Learning, 38(1), 237–257. https://doi.org/10.1111/jcal.12610

Hussain, S. A., Zia, M. H., Zaki, N., & Iqbal, M. Z. (2025). The impact of AI-based learning tools on student motivation and academic self-concept. The Critical Review of Social Sciences Studies, 3(3), 1570–1584. https://doi.org/10.59075/vkgapf12

Kiet Ho Si Thang (2018). The effect of learner autonomy on English proficiency of non-English major students. Journal of Science and Technology Issue on Information and Communications Technology. https://doi.org/10.31130/JST-UD2018-331E

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. RELC Journal, 54(2), 537–550. https://doi.org/10.1177/00336882231162868

Le, H. T. Q., & Nguyen, T. H. (2022). A study on Non-English Major Students' Learner Autonomy: Difficulties and Solutions. International Journal of TESOL & Education, 2(3), 197–207. https://doi.org/10.54855/ijte.222313

Liu, H. J. (2015). Learner Autonomy: The Role of Motivation in Foreign Language Learning. Journal of Language Teaching and Research, 6(6), 1165-1174.

Miao, F., Holmes, W., Huang, R., & Zhang, H. (Eds.). (2021). AI and education: Guidance for policy-makers. UNESCO. https://doi.org/10.54675/PCSP7350

Mohebbi, A. (2024). Enabling learner independence and self-regulation in language education using AI tools: A systematic review. Cogent Education, 12(1), Article 2433814. https://doi.org/10.1080/2331186X.2024.2433814

Nguyen, C. T. N., & Nguyen, T. T. (2020). Non-language-major students' autonomy in learning English in Vietnam. Arab World English Journal, 11(3), 184–192. https://doi.org/10.24093/awej/vol11no3.11

Nguyen, H. U. C. (2024). Factors affecting non-English major students' English proficiency at Hoa Binh University: Current situation and solutions. Hoa Binh University. https://daihochoabinh.edu.vn/factors-affecting-non-english-major-students-english-proficiency-at-hoa-binh-university-current-situation-and-solutions

Nguyen, T. L., & Pham, T. H. (2026). Challenges faced by non-major students in learning English to achieve the outcomes at a Vietnam public university. VNU Journal of Science: Education Research, 41(4). https://doi.org/10.25073/2588-1159/vnuer.4978

Pham, V. T. T., & Pham, A. T. (2025). Investigating EFL students' attitudes towards AI-powered tools in English language learning. 2025 13th International Conference on Information and Education Technology (ICIET), 192–196. https://doi.org/10.1109/ICIET66371.2025.11046282

Pikhart, M., & Al-Obaydi, L. H. (2025). Reporting the potential risk of using AI in higher Education: Subjective perspectives of educators. Computers in Human Behavior Reports, 18, 100693. https://doi.org/10.1016/j.chbr.2025.100693

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Selvi, B., Aldım, Ü. F., & Çoban, Z. (2025). A study on artificial intelligence assisted language learning tools. Tarih Okulu Dergisi, 18(75), 970–999. https://doi.org/10.29228/joh.79085

Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. In K. Hyland & F. Hyland (Eds.), Feedback in second language writing (pp. 125–142). Cambridge University Press. https://doi.org/10.1017/9781108635547.009

Tian, J., & Zhang, R. (2025). Learners' AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy. https://doi.org/10.1016/j.actpsy.2025.105725

Wang, H., Fu, F., Li, L., & Liu, M. (2026). How and when does artificial intelligence use enhance student creativity? The mediating role of knowledge-based dynamic capabilities. Frontiers in Psychology, 17, Article 1861541. https://doi.org/10.3389/fpsyg.2026.1861541

Yang, L. (2026). Empowering the autonomous learner: How AI-assisted language learning environments shape self-regulation, autonomy, and self-directed behaviors. Language Teaching Research. https://doi.org/10.1177/13621688261422129

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of overreliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11(1), Article 28. https://doi.org/10.1186/s40561-024-00316-7

#83 UNPACKING THE BARRIERS: KEY DETERMINANTS OF PUBLIC SPEAKING ANXIETY AMONG ENGLISH MAJOR UNDERGRADUATES AT HUFLIT

Vo Quoc Viet[1]

Abstract

The impact of public speaking anxiety (PSA) was found to significantly affect the quality of presentation among undergraduates majoring in English. This study aims to investigate environmental, psychological, linguistic, and experiential barriers contributing to PSA among English majors at Ho Chi Minh University of Foreign Languages and Information Technology (HUFLIT). This study utilized a convergent parallel mixed-methods, cross-sectional design. Quantitative data collection was carried out via Likert-scale surveys among 133 participants. Results were further established via in-depth semi-structured interviews with 15 highly anxious participants. Results revealed that environmental factors, particularly time pressure and unfamiliar topics (mean=3.85), are the most important statistical stressors. Moreover, the experience of large audiences further heightened anxiety among participants. From a psychological point of view, English majors face an "English major burden" which demands flawless performance (mean=3.79). Such a situation triggers considerable evaluation apprehension. From a linguistic point of view, grammar and vocabulary-related issues evoke considerable anxiety among the participants. Moreover, insufficient practice and rote memorization create an experiential avoidance pattern. The results suggest that these barriers are interconnected. Linguistic hindrances heighten psychological fears, which are further heightened by environmental triggers to create an experiential avoidance pattern. To alleviate PSA among English majors at HUFLIT, it is recommended that participants be desensitized to environmental triggers and encouraged to focus more on effective communication rather than grammar and vocabulary. Additionally, participants should be encouraged to engage in low-stakes practice to overcome rote memorization.

Keywords: public-speaking, anxiety, barriers, mixed-methods

1. Introduction

English major students' target language is not only the tool to communicate with the world, but it is also their main professional orientation. Unfortunately, when students need to present their knowledge in a public performance, the language abilities that once boosted their spirits seem to trigger their fears. Public speaking anxiety (PSA) transforms academic presentations into huge obstacles, especially when the academic and professional identities of those whose job relies heavily on their English proficiency turn into barriers. Therefore, PSA becomes an important challenge that obstructs undergraduates from obtaining academic success and a rewarding career.

[1] *Corresponding author : Vo Quoc Viet, MA, Faculty of Foreign Languages, Ho Chi Minh University of Foreign Languages and Information Technology, Ho Chi Minh city, Vietnam; Email: vietvq@huflit.edu.vn*

1.1. Research Context

For HUFLIT students whose majors are English, PSA appears to be an increasing problem in teaching that needs urgent and thorough examination. As far as observed, a remarkable percentage of English major undergraduates enrolled in public speaking courses in the second semester of 2025 - 2026 have shown medium to high PSA levels. This initial assessment was later confirmed by quantitative data collected with the Personal Report of Public Speaking Anxiety (PRPSA), developed by McCroskey (1970). That specific affective barrier acts as a significant block for students in giving an oral performance in public contexts.

1.2. Research Question

This study concentrates on addressing the following research question: What are the primary environmental, psychological, linguistic, and experiential barriers contributing to public speaking anxiety among English major undergraduates at HUFLIT?

1.3. Purpose of the Study

The present study aims to investigate the contributing factors to PSA in English major undergraduates at HUFLIT. In order to achieve this goal, scientific research will focus on identifying these barriers and proposing evidence for pedagogical approaches aimed at minimizing students' anxiety in public speaking and improving oral presentation confidence.

1.4. Structure of the Paper

The rest of this research paper is organized into five key sections. Firstly, the literature review provides an overview of the previous literature regarding PSA and emphasizes such barriers as psychological, linguistic, environmental, and experiential ones. Secondly, the methodology section includes details about the methods, participants, and procedures used to gather and analyze data related to public speaking anxiety (PSA) among undergraduates at HUFLIT. Thirdly, the empirical data obtained are presented in the Results part, where the crucial factors that determine the emergence of anxiety in the sample are disclosed. Their analysis and interpretation are carried out in the Discussion part, where the obtained results are discussed in the light of existing literature. Finally, the Conclusion and Recommendations part provides the main conclusions and recommendations for English teachers, HUFLIT administration, and all higher education institutions in Vietnam on how to help their students cope with PSA.

2. Literature Review

Public speaking anxiety (PSA) remains a significant challenge to communicative competence and academic success in higher education, especially for the second and foreign language learners. For English majors, public speaking ability is not just an outcome or a requirement but a part of their personal and professional identity, which is why PSA can severely affect both oral production and psychological well-being. For the comprehensive exploration of the complex phenomenon of PSA, this literature review will critically assess recent research across four conceptual dimensions: environmental, psychological, linguistic, and experiential barriers.

2.1. Primary Determinants of Public Speaking Anxiety

2.1.1. Environmental Barriers

The learning environment plays a significant role in intensifying PSA among students. One of the main sources of such anxiety is evaluation anxiety, where students fear negative

evaluation from the instructor and classmates. The presence of external pressure, such as the threat of overt criticism in the classroom, creates a frightening atmosphere for presenters, and instructional and classroom management techniques have a direct impact on the level of stress experienced by the students (Valdivieso Quinaluisa, 2024; Monib, 2025). Strict feedback from the teacher and a lack of emotional and academic support increase student stress. Consequently, when students are expected to perform or respond orally without the foundation of a psychologically safe, communicative environment, they inevitably experience debilitating performance pressure.

2.1.2. Psychological Barriers

Extensive research demonstrates that students' PSA are affected by internal psychological processes. In particular, Le (2024) and Qayyum et al. (2024) found that intrinsic features, especially a student's overall confidence in learning a second or foreign language, contribute significantly to the level of speaking anxiety. Such anxiety is often expressed as strong psychological barriers. For example, Boonchuayrod (2025) revealed that students had high cognitive anxiety in answering spontaneous questions during public speaking. Moreover, Qayyum et al. (2024) pointed out that the fear of making mistakes and the lack of confidence to speak frequently result in students' silence and hinder their oral production. From a motivational perspective, Alamer and Almulhim (2021) pointed out that being competent and socially connected aided in reducing particular kinds of anxieties. In contrast, controlled motivation heightened overall language anxiety.

2.1.3. Linguistic Barriers

Language anxiety is associated with language learners' assumed communicative competence; a student who assumes their competence in the target language is low is more anxious than those who assume high competence in a communicative situation. According to Qayyum et al. (2024), these language constraints are identified as the major hindrances in communication. Moreover, English-speaking ability is identified as a vital aspect influencing PSA at all stages of performance (Viwattanabunchong, 2017).

From the micro-linguistic point of view, specific deficiencies in vocabulary and grammar hinder performance in speaking (Monib, 2025; Qayyum et al., 2024). Monib (2025) found that excessive concern for perfect grammar while speaking spontaneously interferes with communication. Besides the restrictions imposed by syntax, phonological precision is equally important; according to Karmida et al. (2024), mispronunciation reduces the student's level of oral competence and hence affects their confidence. Therefore, as illustrated by Le (2024), a lack of knowledge or confidence in one's linguistic ability causes emotional distress because insufficient vocabulary and inaccurate pronunciation create intense anxiety in academic presentations.

2.1.4. Experiential Barriers

The intensity of PSA varies depending on the specific stage of the assignment, the academic level of a student, and institutional curriculum design. Viwattanabunchong (2017) described PSA in four separate stages of task performance: pre-preparation, preparation, pre-performance, and performance. Whereas research shows that anxiety persists through all four stages, it peaks at the performance stage. However, the intensity of anxiety that arises at this final stage depends heavily on the field of study of the student and their educational level. To be specific, English majors report significantly lower performance anxiety than non-majors, owing to a curriculum designed in such a way that there is regular oral production practice and continuous formative instructions (Boonchuayrod, 2025). Additionally, the degree of education

is an important factor when determining the intensity of PSA. Undergraduates and students in master's programs have twice as much anxiety as doctoral students (Lintner & Belovecov, 2024).

2.2. Research Gap

Despite a large number of previous studies focusing on PSA in a global context (e.g., Thailand, South Korea, Brunei, Saudi Arabia, Indonesia, Ecuador, the Czech Republic, the Philippines, and Pakistan), there is still a gap in the existing literature. Specifically, there has been a lack of empirical research investigating the combined effects of environmental, psychological, linguistic and experiential barriers on PSA of English major undergraduates in Vietnam higher education context.

2.3. Rationale for Conducting the Research at HUFLIT

An investigation into the PSA of English major undergraduates at HUFLIT is justified by the following institutional and instructional considerations:

2.3.1. Pedagogical Optimization and Curricular Refinement

Frequent oral practice often serves as an effective remedy to alleviate PSA, yet the emphasis on stringent evaluation and competition might create a paradoxical pedagogical situation wherein practice, by increasing stress levels, could exacerbate PSA rather than alleviate it. The rationale behind this issue is: If practice causes a considerable amount of anxiety in the classroom, the educational interventions currently in use may be less effective in achieving the desired impact on performance outcomes. Research findings have indicated that the alleviation of PSA needs instructional methods that can create a comfortable, student-centered environment, and consequently, identifying the particular aspects of environmental, psychological, and linguistic and experience based that lead to the emergence of PSA among HUFLIT English major undergraduates will permit us to propose and to implement appropriate pedagogical interventions for mitigating such negative affect. In addition, the research outcomes will provide empirical evidence for effective evaluation and modification of the present speaking curricula.

2.3.2. Alignment with Academic Standards and Professional Expectations

HUFLIT is a higher education institution specializing in training students in foreign languages and international communication. Its mission involves the development of highly proficient linguistic graduates to prepare them for direct entry into the global labor market. This challenging goal has resulted in a competitive, high-stakes learning environment that can potentially induce significant distress among undergraduates. In such an environment, PSA is viewed as a substantial hindrance to the successful academic performance and later employability of the graduates. Hence, considering the training objectives of the institution for future international communicators, PSA represents a challenge for its fulfillment and also a pedagogical impediment for effective teaching.

3. Methodology

This section will describe the methods of study that were applied to reveal the factors contributing to PSA amongst the English major undergraduates of HUFLIT. It gives comprehensive information about the research design, participants, sampling strategy, tools used to gather data and methods of analyzing data.

3.1. Research Design

A mixed-method cross-sectional design was adopted in this research. The combination of different methodologies allowed not only to broaden but also to deepen the obtained data. In particular, a convergent parallel design was applied, whereby quantitative and qualitative data were collected at the same time, analyzed separately and then merged at the interpretation stage.

3.2. Population and Sample

The targeted population for this study included undergraduate students majoring in English at HUFLIT who either currently took or had finished attending the public speaking class. The following dual-sampling design was employed to select participants for the research:
3.2.1. Quantitative Component:

A convenience sampling technique was used to recruit 133 undergraduate students who would take part in completing the main survey.

3.2.2. Qualitative Component:

A purposeful sampling technique was used to choose a particular subset of 15 highly-anxious students out of all participants recruited for conducting qualitative semi-structured interviews.

3.3. Data Collection Instruments

Data collection involved two tools that were developed to satisfy mixed-method research requirements:

3.3.1. Online Survey:

The questionnaire, based on a Likert scale consisting of 5-point items, aimed at gathering data about the broadness of information belonging to the obstacles contributing to PSA: environmental, psychological, linguistic, and experiential barriers.

3.3.2. Semi-structured Interviews:

In order to find appropriate participants for the second component of the study, the Personal Report of Public Speaking Anxiety (PRPSA) was conducted among all selected participants. Subsequently, fifteen students with the highest scores on the PRPSA test were interviewed. The interviews used open-ended questions that would enable participants to reflect upon their experiences and identify internal barriers.

3.4. Data Collection Procedure

All the processes involved in the data collection were carried out systematically over six weeks in January and February 2026 as follows:

3.4.1. Ethics and Consent:

Ethical approval was granted, and informed consent was digitally gathered from all 133 participants before collecting the required data. Participants' rights, privacy, and the objective of the study were clearly informed.

3.4.2.Survey Distribution:

The survey on determinants of PSA and PRPSA was conducted simultaneously using Google Forms.

3.4.3. Interview Execution:

After 15 participants with the highest scores in PRPSA had been recognized, they were interviewed utilizing the semi-structured interview format. Each of these interviews took about 20 to 30 minutes to complete. All interviews were recorded after obtaining verbal consent from each participant.

3.5. Data Analysis

Both quantitative and qualitative data analyses were carried out separately before being triangulated and merged into one analysis.

3.5.1. Quantitative Analysis:

Descriptive analysis was performed on the data obtained through the Likert scale survey to measure the average scores of survey items and standard deviation values.

3.5.2. Qualitative Analysis:

The interview transcripts were analyzed using thematic analysis established by Braun and Clarke (2006). This involved the coding process for generating themes related to identifying environmental, psychological, linguistic, and experiential barriers contributing to PSA based on participants' personal perceptions and experience.

3.5.3. Integration:

As a result of adopting the convergent parallel design in this study, the descriptive statistics results and themes from the qualitative data were compared in order to identify the convergence or divergence between the two data sources related to the research question.

3.6. Expected Outcomes and Limitations

The first and foremost expected outcome is the identification of environmental, psychological, linguistic, and experiential barriers that contribute to PSA among HUFLIT English major undergraduates and provide insights for pedagogical interventions. Nonetheless, certain limitations must be admitted in regard to this research methodology. First, the use of surveys and interviews may lead to methodological limitations in the research, because of respondent self-reporting bias, in which participants may over- or under-report their degree of anxiety, based on social norms or personal subjective memory. Second, the generalizability of the findings to English majors at other universities is potentially limited, due to the nature of the sample drawn from one site (HUFLIT).

4. Results

A convergent parallel mixed-methods design was utilized to integrate the quantitative and qualitative datasets. This triangulation facilitates a comprehensive, holistic understanding of the primary determinants contributing to PSA among English major undergraduates at HUFLIT.

4.1 Environmental Barriers

As illustrated in Table 4.1, environmental factors emerged as the most significant barrier, yielding the highest overall mean score among all the analyzed categories (overall mean=3.75). The factors with the greatest impact on students' performance include the apprehension toward unfamiliar topics alongside time pressure (mean=3.85), closely followed by presentations in front of a large audience (mean=3.80). Correspondingly, interviewees noted that public speaking in front of a big crowd deepens the feeling of isolation and vulnerability, while heightening a sense that the audience is uninterested. Moreover, strict time pressure intensifies students' anxiety levels.

Table 4.1

Environmental Determinants Contributing to PSA

Items	N	Mean	Std Deviation	Min	Max
Large audiences increase my speaking anxiety.	133	3.796992	1.172665	1	5
Strict lecturer evaluation heightens my nervousness.	133	3.676692	1.11157	1	5
Peer pressure or comparison with classmates affects me.	133	3.669173	1.112748	1	5
Unfamiliar topics or time limits make me anxious.	133	3.849624	1.048059	1	5
Overall mean:		3.74812			

4.2. Psychological Barriers:

Table 4.2 demonstrates that psychological barriers are the second most impactful obstacle in causing PSA (overall mean=3.67). Primarily, this result is driven by intense perfectionist tendencies within the students' major field of study (mean=3.79), a concern highlighted during their discussion of the "Burden of the English Major Identity", where students felt a heavy institutional and professional pressure to use language perfectly.

Table 4.2

Psychological Determinants Contributing to PSA

Items	N	Mean	Std Deviation	Min	Max
I worry about others judging or criticizing my speech.	133	3.736842	1.05812	1	5
I lack confidence in my ability to deliver a good speech.	133	3.654135	1.08044	1	5
I fear forgetting my content or making mistakes in front of an audience.	133	3.781955	1.032343	1	5
I avoid voluntary speaking activities due to fear.	133	3.406015	1.206282	1	5
I feel pressure to perform perfectly due to being an English major.	133	3.789474	0.992797	1	5
Overall mean:		3.673684			

4.3. Linguistic Barriers:

As indicated in the quantitative data in Table 4.3, linguistic barriers were found to be a major problem for students. The mean score for the overall category is 3.64, which indicates a high level of anxiety in their competence in the target language. The greatest obstacle in the category for PSA was limited vocabulary, with a mean score of 3.80, closely followed by concerns about accurate grammar, with a mean score of 3.68. Supporting the quantitative results, qualitative interviews revealed how the fear of grammatical errors and the possibility of being misunderstood put non-native English major presenters in a state of cognitive self-monitoring. The constant evaluation of the grammatical structure of each utterance overloads their cognitive capacity, so that they have to prioritize the correctness of the syntax over the natural flow of communication.

Table 4.3

Linguistic Determinants Contributing to PSA

Items	N	Mean	Std Deviation	Min	Max
My limited vocabulary causes anxiety during speeches	133	3.796992	1.099307	1	5
I struggle with pronunciation/fluency in English presentations.	133	3.413534	1.194227	1	5
Grammar errors make me nervous when speaking publicly.	133	3.684211	1.032564	1	5
I feel my English skills are not sufficient for academic presentations.	133	3.676692	1.171304	1	5
Overall mean:		3.642857			

4.4. Experiential Barriers:

Experiential barriers had one of the lowest overall mean scores in the quantitative sections (mean = 3.45), and the range of responses is shown in Table 4.4. The highest mean in the category is lack of practice (mean=3.75). Even though previous negative experiences scored lower (mean = 3.25). While past negative experiences had a lower average score of 3.25, qualitative findings indicate that these experiences still deeply affect undergraduate students' PSA. These experiences can create a negative internal dialogue that leads students to self-handicap and to actively work against their academic success. Besides this negative internal dialogue, participants noted that avoidance of speaking opportunities, often due to past negative experiences, leads directly to a lack of practice. Thus, students don't get the systematic desensitization that's necessary to lower their anxiety, and they also miss out on opportunities to reframe their prior negative experiences more positively. Another way students try to overcome their fear of forgetting their content is through rote memorization. This coping strategy is ineffective because the fixed, linear nature of memorized text means that forgetting a single transitional word can cause the organization of the entire presentation to collapse.

Table 4.4

Experiential Determinants Contributing to PSA

Items	N	Mean	Std Deviation	Min	Max
Lack of practice opportunities contributes to my anxiety.	133	3.75188	1.03284	1	5
Negative past feedback from lecturers/peers affects my speaking confidence.	133	3.24812	1.227254	1	5
Previous bad experiences (e.g., embarrassment) linger.	133	3.315789	1.1894	1	5
Insufficient feedback/training worsens my anxiety.	133	3.466165	1.062793	1	5
Overall mean:		3.445489			

5. Discussion

The primary objective of this research was to investigate the environmental, psychological, linguistic, and experiential barriers contributing to PSA among English major

undergraduates at HUFLIT. The integration of the quantitative and qualitative analyses reveals that these four types of barriers are not separate but interconnected to form a loop of anxiety. The environmental factors have the strongest impact on students' PSA, followed by psychological and linguistic barriers, while experiential factors show the lowest anxiety levels, but cause long-term avoidance of public speaking practice.

5.1. Environmental Barriers: Unfamiliar Topics, Time Pressure, and Audience Scale

Among various determinants, environmental barriers recorded the highest mean score for PSA among HUFLIT English major undergraduates (overall mean = 3.75), with particular contributions from unfamiliar topics, strict time pressure (mean = 3.85) and audience scale (mean = 3.80). These results partially align with the studies of Valdivieso Quinaluisa (2024) and Monib (2025), who noted that external pressure, such as being ridiculed or criticized by other people, creates an unfriendly and threatening environment for students to perform ineffectively. However, this research extends beyond these foundational works because it shifts the main concentration of environmental anxiety. Although previous literature tends to focus on the effects of the environmental pressure of comparing oneself to peers and being judged by teachers, HUFLIT undergraduates found that the situational limitations of the testing environment, such as a large audience and strict time pressure, created a strong feeling of isolation. This implies that the situation - performing in front of many people under strict time pressure - has a more noticeable influence on the occurrence of PSA than peer comparisons.

5.2. Psychological Barriers: "The Burden of the English Major Identity"

Psychological barriers emerged as the second-most influential factor contributing to PSA (overall mean = 3.67). This is consistent with earlier studies conducted by Le (2024) and Qayyum et al. (2024) that point out low self-confidence levels and cognitive fears as the main causes for speech anxiety. The results obtained from this study reinforce their theory, particularly regarding the fear of forgetting content or making mistakes (mean = 3.78). Crucially, our results add a new perspective to this psychological aspect in terms of "The Burden of the English Major Identity" (mean = 3.79). Unlike general language learners, HUFLIT English majors face intense, internalized pressure for perfectionism rooted in their chosen specialization. The result from this study directly contradicts the theory of Alamer & Almulhim (2021), which proposes that competence and relatedness mitigate specific forms of anxiety. For the students in the study, public speaking is closely connected to their academic and professional identity, transforming normal presentations into high-stakes evaluations. This intense pressure leads to cognitive anxiety and fear of evaluation.

5.3. Linguistic Barriers: The Cognitive Self-Monitoring State

Linguistic barriers are major contributors to students' PSA, with an overall mean of 3.64. The main concerns were vocabulary limitations (mean = 3.80) and the fear of grammatical inaccuracy (mean = 3.68). The findings of this study support the theories advanced by Qayyum et al. (2024) and Viwattanabunchong (2017) that language competence is essential in developing speaking confidence. Moreover, the qualitative integration provides a profound insight into the psychological aspect of the theory about grammar discussed by Monib (2025), who states that overemphasizing grammar impedes fluency. Empirical evidence of this issue can be found in the current study in the form of the cognitive self-monitoring state. Under the fear of being misunderstood or negatively evaluated, presenters become locked in an ongoing process of real-time editing to ensure syntactic accuracy. Consequently, this intense cognitive

load and psychological strain result in anxiety actions like hand shaking and trembling of the voice.

5.4. Experiential Barriers: The Avoidance of Public Speaking Practice and Rote Memorization

Experiential barriers got the lowest overall quantitative average score (mean=3.45). This finding supports Boonchuayrod's (2025) argument that English majors have less performance anxiety than non-majors due to an immersive curriculum that includes a lot of presentations. However, qualitative data deeply undermine the assumption that low frequency or low quantitative scores imply a low psychological impact. For instance, although a lack of practice opportunities scored relatively high within this category (mean = 3.75), negative past feedback scored the lowest numerically (mean = 3.25), yet it serves as a powerful psychological trigger. Rather than fading over time, these negative past experiences form destructive internal dialogues that result in "self-handicapping," which ultimately drives students to avoid practice. By avoiding voluntary speaking opportunities, students prevent themselves from gaining the practice needed to reframe those negative memories into positive experiences. Students tend to rely on rote memorization to avoid the anxiety of repeating past failures. In theory, this coping strategy is precarious; the loss of one connecting word can bring down the entire edifice of a presentation, ultimately resulting in a poor performance.

5.5. The Loop of Public Speaking Anxiety

This paper provides comprehensive insights into PSA, demonstrating how four types of determinants are interrelated. First, language insecurity is the starting point of this process, where the need to be grammatically perfect, syntactically controlled, and to use a limited vocabulary creates a strong fear of being evaluated negatively. This fear of negative evaluation psychologically is intensified by the environmental factors, such as large audience size, high institutional expectations and time pressure. All these determinants together lead to a high level of performance stress, which makes students adopt ineffective coping strategies such as avoiding public speaking practice and using rote memorization. These strategies do not bring relief or improve linguistic competence, but they actually make the original language anxiety worse, and keep the loop going.

6. Conclusion

Employing a convergent parallel design to examine the key determinants of PSA of HUFLIT English major undergraduate students, the results of this study confirm that PSA is caused by a network of interconnected environmental, psychological, linguistic, and experiential obstacles. Specifically, environmental barriers were found to be the most severe constraint (overall mean=3.75), mainly due to limited time to present unfamiliar topics, and the feeling of isolation when speaking in front of a large crowd. Closely following this, psychological barriers ranked second (overall mean=3.67), because of students' burden of being an English major, who had to speak without making mistakes, in addition to the fear of being evaluated. Similarly, linguistic barriers (overall mean=3.64) rooted in insufficient vocabulary and grammar fear, which led to an intense state of self-monitoring in which students over-edit their sentences during performance, thus losing focus on effective communication. On the other hand, experiential barriers had the lowest mean score (overall mean=3.45). Past negative learning experiences and feedback, along with the psychological and linguistic demands for perfect delivery in front of an audience, form an invisible barrier that compels students away from practicing public speaking, which leads to insufficient preparation, thus greatly triggering PSA.

Recommendations

Based on the findings of this study, the following pedagogical and institutional recommendations are proposed to mitigate PSA among students.

For Educators

- *Scaffold Presentation Difficulty:* Students should be allowed to choose familiar topics for their initial major presentations. This scaffolding will enable students to build confidence and develop time management skills before engaging with more complex subject matters.
- *Foster collaborative Interaction:* To mitigate the sense of isolation experienced by the presenter, interactive components such as structured question and answer sessions should be incorporated in lesson planning. This design shifts frightening one-way presentations into collaborative conversations.
- *Reorient Evaluation Rubrics:* Communicative competence and presentation efficacy should be prioritized over grammatical accuracy in assessment criteria. By allowing students to speak without disciplinary measures for minor mistakes, the real-time grammar self-correction process will be broken. This will allow for a more natural delivery.
- *Discourage Rote Memorization:* Instructors should avoid requiring students to deliver perfect content that is based on memorization during their presentations. Instead, build confidence in their speaking and give them practical tools to present well through a short outline using improvised practice.

For the University:

- *Establish Low-Anxiety Practice Environment:* Universities should create non-threatening conditions for students to practice public speaking without being graded, through review sessions, English speaking clubs, or debating contests, where students can accumulate positive experience in making presentations and alleviate anxiety
- *Build Positive Communicator Identity:* Provide institutional orientation or counseling programs so that the burden of linguistic perfection expected from non-native English major students can be eliminated. Students should realize that making errors is inevitable in language learning.

References

Alamer, A., & Almulhim, F. (2021). The interrelation between language anxiety and self-determined motivation: A mixed methods approach. *Frontiers in Education*, *6*, Article 618655. https://doi.org/10.3389/feduc.2021.618655

Boonchuayrod, P. (2025). Public speaking anxiety in higher education: A comparative analysis of English-major and non-English-major undergraduates. *International Journal of Development Administration Research*, *8*(2), 153–172.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Karmida, K., Wariyati, W., Hasibuan, A. L., & Nurlaili, N. (2024). An analysis of students' low confidence in speaking, as the attribute of pronunciation context. Journey: Journal of English Language and Pedagogy, 7(2), 331–336. https://doi.org/10.33503/journey.v7i2.845

Le, Y. (2024). Factors causing college students' English speaking anxiety. *Applied & Educational Psychology*, *5*(1), 159–164. https://doi.org/10.23977/appep.2024.050122

Lintner, T., & Belovecová, B. (2024). Demographic predictors of public speaking anxiety among university students. *Current Psychology*, *43*(30), 25215–25223. https://doi.org/10.1007/s12144-024-06216-w

McCroskey, J. C. (1970). Measures of communication-bound anxiety. *Speech Monographs*, *37*(4), 269–277. https://doi.org/10.1080/03637757009375677

Monib, W. K. (2025). Determinant factors influencing English speaking skill among undergraduates: Challenges and remedial strategies. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, *10*(1), 72–85. https://doi.org/10.35974/acuity.v10i1.3351

Qayyum, N. U., Zahoor, N. I., & Bibi, N. M. (2024). Breaking the silence: Unveiling the hidden challenges to English speaking among undergraduates. *Voyage Journal of Educational Studies*, *4*(3), 197–210. https://doi.org/10.58622/vjes.v4i3.197

Valdivieso Quinaluisa, M. G. (2024). *Speaking anxiety in oral production of the English major undergraduate students* (Undergraduate research report, Technical University of Cotopaxi, Pujilí Campus).

Viwattanabunchong, S. (2017). *A survey of factors that cause public speaking anxiety for EFL learners* [Master's thesis, Thammasat University]. Language Institute, Thammasat University.

Appendix

Links to access data collection instruments and datasets

1. Quantitative survey (statements) :

https://docs.google.com/document/d/1TFabSWxGaboPr_JMATcK4kSf1KieKb74/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

2. Quantitative survey (responses) :

https://docs.google.com/spreadsheets/d/14GUvC8PSskI5Jlty8qYP6HJ9ntk9MyFD/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

3. Qualitative semi-structured interview questions:

https://docs.google.com/document/d/1_rIrX_U4rSd-ChoQtHEjToKRu95hZzdC/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

4. Qualitative semi-structured interview (responses):

https://docs.google.com/spreadsheets/d/144MSojPKwBVn5YmEIIceMCujnI_Qxmsh/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

5. The Personal Report of Public Speaking Anxiety (PRPSA)-Introversion Scale (statements):

https://docs.google.com/document/d/1drwqnVco5qZ3BTDHHI0Fnob_LUnpQ1Ev/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

6. The Personal Report of Public Speaking Anxiety (PRPSA)-Introversion Scale (responses):

https://docs.google.com/spreadsheets/d/19xrru1cKjMLD6BZhKrK2h3Znjz6_49PP/edit?usp=sharing&ouid=106944625484452431759&rtpof=true&sd=true

#84 ESL INTEGRATION IN JAPANESE READING INSTRUCTION: EXPLORING THE ROLE OF ENGLISH IN SUPPORTING JAPANESE READING COMPREHENSION AMONG SENIOR JAPANESE-MAJOR STUDENTS AT VNU-HCM UNIVERSITY OF SOCIAL SCIENCES AND HUMANITIES

Vu Doan Lien Khe[1]

Abstract

As Vietnam transitions toward an ESL environment, many students face "language attrition" where their English proficiency declines while focusing on a third language (L3) like Japanese. This case study explores a practical pedagogical approach implemented in Japanese reading classes for fourth-year students at the University of Social Sciences and Humanities (USSH). To tackle this issue, the author introduced a bilingual dynamic that replaces traditional Japanese-Vietnamese translation with Japanese-English "live translation" and vocabulary mapping.

This method encourages students to bridge two foreign languages, gradually creating a cognitive link that may strengthen both. By mapping complex Japanese structures to English counterparts, students deepen their understanding of Japanese nuances while keeping their English active and functional for professional use. To evaluate the model, a mixed-methods research design was employed, involving surveys and qualitative feedback from senior Japanese majors. To make classroom observations more systematic, students' "live translation" performance was observed in class with a focus on semantic accuracy, lexical equivalence, and response latency.

Findings suggest that ESL integration can enhance Japanese textual analysis and is associated with higher reported confidence and cross-linguistic flexibility for many students. Such results are vital for graduates entering a competitive labor market where "English+1" is a highly sought-after asset. This study offers a practical framework for educators to create English-rich environments within diverse curricula, aligning with the HUFLIT INTESOL 2026 goals of fostering English as a versatile tool for student employability in Vietnam's evolving education landscape.

Keywords: ESL integration, Japanese reading, L3 acquisition, English attrition, HCMUSSH, multilingualism.

1. Introduction

In the current educational landscape of Vietnam, the transition toward an ESL (English as a Second Language) environment has placed unique demands on multilingual learners. At the Department of Japanese Studies, Faculty of Oriental Studies, University of Social Sciences and Humanities, VNU-HCM, senior students often face a subtle but persistent challenge:

[1] *Ph. D., University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City, Faculty of Oriental Studies* vukhejp@hcmussh.edu.vn

“language attrition”. While they are heavily focused on mastering Japanese (L3) for their future careers, their English (L2) skills, which they spent years building, tend to stagnate or “fade away” due to a lack of practical use in the specialized curriculum.

For years, the standard approach to teaching Japanese reading has been a direct bridge between Japanese and Vietnamese. While this helps with basic comprehension, it unintentionally creates a linguistic wall that keeps English out of the picture. In today’s job market, where "English + 1" is often the minimum requirement, letting English proficiency slip away during the final years of university represents a missed opportunity for graduates.

The pedagogical approach investigated in this study emerged from the author’s classroom observations while teaching fourth-year Japanese reading courses. Many students demonstrated strong Japanese reading skills yet struggled to maintain confidence in English communication. To address this issue, the author introduced a bilingual instructional model in which students were encouraged to use English as a cognitive bridge when interpreting Japanese vocabulary, grammar, and textual meanings. Rather than relying exclusively on Japanese-Vietnamese translation, students were invited to explore Japanese-English equivalences and participate in spontaneous “live translation” activities. The intention was not to replace Vietnamese but to create additional opportunities for students to activate English in a meaningful academic context.

Despite growing interest in multilingual education, translanguaging, and cross-linguistic transfer, relatively little research has examined the integration of English into Japanese reading instruction within Vietnamese higher education. Previous studies have primarily focused on bilingual education or the acquisition of individual foreign languages, while the pedagogical potential of English as a supporting language for Japanese reading comprehension remains underexplored. Consequently, empirical evidence regarding students’ perceptions, benefits, and challenges associated with this instructional approach is still limited.

Building upon theories of language attrition, cross-linguistic transfer, and multilingual education, this study aims to examine whether the integration of English into Japanese reading instruction can serve as an effective instructional approach for maintaining students’ English proficiency while supporting Japanese language development. More specifically, the study investigates students’ awareness of English attrition, their perceptions of the bilingual instructional model, the influence of English proficiency on those perceptions, and the advantages and challenges associated with using English as a mediating language in Japanese reading classes. In doing so, the study contributes empirical evidence to ongoing discussions surrounding multilingual pedagogical practices in Vietnamese higher education.

Research Objectives

Based on the aforementioned research gap, this study aims to examine the effectiveness of integrating English into Japanese reading instruction as a pedagogical strategy for supporting multilingual language development. More specifically, this study pursues the following objectives: (1) to investigate students’ awareness of English attrition, (2) explore their perceptions of the bilingual instructional model, (3) examine whether these perceptions vary according to English proficiency levels, and (4) identify the benefits and challenges associated with the use of English as a mediating language in Japanese reading classes.

Research Hypotheses

Guided by the theoretical perspectives of language attrition, cross-linguistic transfer, translanguaging, and cognitive load, the study proposes the following hypotheses:

H1: Students will generally demonstrate positive perceptions toward the integration of English into Japanese reading instruction.

H2: Students with higher levels of English proficiency will report more positive perceptions of the bilingual instructional model than students with lower levels of proficiency.

H3: The integration of English into Japanese reading instruction can contribute to maintaining students' engagement with English while facilitating Japanese reading comprehension.

2. Theoretical framework and methodology

2.1 Theoretical framework

2.1.1. Language attrition in the multilingual learning contexts

This study begins with the concept of *Language Attrition*, which is the gradual fading of a learner's previously acquired English (L2) as their attention shifts toward mastering Japanese (L3) (Kopke & Schmid, 2004). In a multilingual setting, language systems continuously interact and compete for mental effort rather than functioning independently (Herdina & Jessner, 2002). This research aims to reactivate learners' underused English by giving it a meaningful role within the Japanese learning environment. Consequently, when learners devote substantial attention to acquiring a new language, previously learned languages may become less accessible over time.

This phenomenon is particularly relevant in the Vietnamese higher education context, where students majoring in Japanese often devote most of their academic efforts to developing Japanese proficiency while using English less frequently. Recent studies suggest that language attrition is not merely a matter of forgetting vocabulary or grammar but also involves reduced confidence, decreased fluency, and diminished willingness to use the language actively (Schmid & Kopke, 2017; 2019). Therefore, maintaining meaningful opportunities for English use throughout university education has become an important pedagogical concern.

Within this context, the present study considers ESL integration as a potential strategy for mitigating language attrition by providing students with opportunities to reactivate English while engaging in Japanese reading activities.

2.1.2. The Interdependence Hypothesis and Cross-linguistic transfer

Rather than viewing languages as rivals, this study draws on Cummins' *Interdependence Hypothesis*, which proposes that linguistic abilities can be shared and reactivated through a *Common Underlying Proficiency* (Cummins, 2000). According to this framework, different languages share a *Common Underlying Proficiency (CUP)*, through which linguistic knowledge, cognitive skills, and literacy abilities may transfer across languages.

This perspective is closely associated with the concept of cross-linguistic transfer, which refers to the influence of one language on the learning and processing of another. Research has demonstrated that multilingual learners frequently draw upon previously acquired linguistic knowledge when interpreting new language forms and meanings (Cummins, 2021). Such transfer may occur at lexical, semantic, syntactic, and metacognitive levels.

By encouraging students to connect Kanji meanings with their English equivalents, the model applies *Translation in Language Teaching (TILT)* not as rote practice, but as a bridge that strengthens both languages in tandem.

In the present study, English is conceptualized as a mediating language that can facilitate Japanese reading comprehension. By encouraging learners to identify connections between Japanese expressions and English equivalents, the instructional model seeks to activate prior linguistic knowledge and strengthen multilingual awareness. Rather than viewing English and Japanese as competing systems, this approach treats them as complementary resources that may mutually reinforce learning.

In the Vietnamese context, the transition towards an ESL environment as part of the National Foreign Languages Project has redefined how students approach multilingualism. Researchers emphasize that for Vietnamese learners, the interaction between multiple foreign languages requires a strategic pedagogical shift to optimize cognitive resources (Nguyen & Nguyen, 2019).

2.1.3. Translanguaging as a multilingual pedagogical practice

Recent developments in multilingual education have challenged traditional assumptions that languages should be kept strictly separate in classroom instruction. Translanguaging theory argues that multilingual learners possess an integrated linguistic repertoire from which they flexibly draw according to communicative needs (Garcia & Wei, 2014; Wei, 2017).

The classroom model examined in this study aligns with translanguaging principles by encouraging students to utilize English when interpreting Japanese vocabulary, grammar, and textual meanings. Such practices enable learners to mobilize their full linguistic resources while developing greater awareness of relationships among languages. Furthermore, translanguaging has been identified as an effective approach for supporting multilingual skills in increasingly globalized educational environments (Cenoz & Gorter, 2022).

2.1.4. Cognitive load theory and the language threshold

Despite the potential benefits of multilingual pedagogies, the simultaneous activation of multiple languages may also increase cognitive demands. *Cognitive Load Theory* proposes that learning effectiveness depends on the limited capacity of working memory to process information (Paas & Ayres, 2014). When instructional tasks require learners to manage several linguistic systems simultaneously, excessive cognitive load may hinder comprehension and learning outcomes.

Related to this concern, the *Language Threshold Hypothesis* suggests that learners require a minimum level of proficiency before they can effectively benefit from cross-linguistic transfer (Cummins, 2017). Students with stronger English proficiency may be better equipped to utilize English as a cognitive support tool when reading Japanese texts, whereas learners with weaker English backgrounds may experience confusion, mental fatigue, or reduced confidence.

Consequently, the effectiveness of ESL integration is likely to vary according to learners' existing language competence. This perspective provides an important framework for interpreting differences in students' perceptions and experiences within the bilingual instructional model investigated in the present study.

2.1.5. Conceptual framework of the study

Drawing upon *Language Attrition Theory*, the *Interdependence Hypothesis*, *Translanguaging Theory*, and *Cognitive Load Theory*, the present study conceptualizes ESL integration as a pedagogical mechanism through which English can function as a cognitive

bridge in Japanese reading instruction. The framework assumes that meaningful use of English may help reduce language attrition and facilitate cross-linguistic transfer; however, the effectiveness of such an approach may be moderated by learners' existing English proficiency and cognitive capacity. These theoretical perspectives collectively inform the investigation of students' perceptions, experiences, and responses to ESL-integrated Japanese reading activities.

2.2 Methodology

This study employed a mixed-methods case study design to capture the implementation of ESL integration in Japanese reading instruction. A mixed-methods approach was selected because it enables the researcher to combine quantitative evidence with qualitative insights, thereby providing a more comprehensive understanding of students' perceptions and experiences regarding the bilingual instructional model.

This research was carried out at the Department of Japanese Studies, Faculty of Oriental Studies, USSH, VNU- HCM. The participants consisted of 51 fourth-year students enrolled in a Japanese Reading Comprehension course during the 2025-2026 academic year. These students were selected because they had accumulated substantial learning experience in both English and Japanese and were therefore in a suitable position to reflect on issues related to language attrition, multilingual learning, and cross-linguistic interaction. Furthermore, as students approaching graduation, they represented a population for whom maintaining both English and Japanese proficiency is particularly relevant to future employability.

The scope of this study was deliberately limited to students' perceptions, attitudes, and experiences regarding the integration of English into Japanese reading activities. The research did not attempt to measure actual improvements in English or Japanese proficiency through standardized language tests. Instead, the study focused on how learners perceived the effectiveness, practicality, benefits, and challenges of the instructional approach. Consequently, the findings should be interpreted within the specific context of Japanese-major students at a single institution and may not be generalized to other educational settings.

Data were collected through a structured survey comprising two complementary parts. The quantitative section used 5-point Likert scale items to measure student satisfaction, vocabulary retention, and perceived cognitive effort. The qualitative section invited open-ended reflections, giving space for students to share their challenges and feelings of mental fatigue. Together, these data sources provided both statistical rigor and a more personal view of the classroom atmosphere. Prior to data collection, the questionnaire was reviewed by two specialists in language education to ensure content validity and alignment with the research objectives. Minor revisions were made based on their feedback to improve clarity and relevance.

In addition to the questionnaire, classroom observations during "live translation" sessions were used to note students' use of English as a mediating language. The observations focused on semantic accuracy, lexical equivalence and response time, there by providing qualitative support for interpreting the survey results.

The internal consistency of the questionnaire was examined using Cronbach's alpha. The overall reliability coefficient for the nine Likert-scale items was $\alpha=.789$, indicating acceptable internal consistency according to commonly accepted standards for educational research (Taber, 2017). Furthermore, the subscale measuring perceptions of ESL integration demonstrated satisfactory reliability ($\alpha=.758$). These results suggest that the instrument was

sufficiently reliable for investigating students' perceptions of English attrition and ESL integration in Japanese reading instruction.

Table 1. Reliability analysis of the questionnaire

Scale	Number of items	Cronbach's Alpha
English attrition	2	.506
ESL integration effectiveness	5	.758
Career relevance	2	.638
Overall questionnaire	9	.789

As shown in **Table 1**, the overall questionnaire achieved an acceptable level of internal consistency (α = .789). The ESL Integration Effectiveness subscale demonstrated the highest reliability (α = .758), while the lower coefficients observed in the two-item subscales may be attributed to the limited number of items rather than substantial measurement problems.

Quantitative data were analyzed using SPSS 26.0. Descriptive statistics (Mean and Standard Deviation) established a baseline for student perceptions, followed by a one-way ANOVA comparing three proficiency groups, Pre-intermediate, Intermediate, and Advanced, to examine whether perceived "cognitive overload" differed significantly among them. Given that the classroom model relies on English as a mediating language, it was considered essential that participants possess at least a pre-intermediate level of L2 competence. Students with very limited English proficiency were not targeted in this case study, as they might be more vulnerable to cognitive overload rather than benefiting from scaffolding effects. Finally, thematic coding was applied to the qualitative responses, allowing the numerical findings to be interpreted through the authentic voices and experiences of the participants.

Qualitative responses were analyzed using thematic analysis following the framework proposed by Braun and Clarke (2006). The analysis involved familiarization with the data, generation of initial codes, identification of recurring themes, review and refinement of themes, and interpretation of findings. Coding was conducted iteratively to enhance consistency, and the resulting themes were continuously compared with the original responses to ensure interpretive validity. The qualitative findings were subsequently used to complement and explain patterns identified in the quantitative analysis.

2.3 Research delimitations and limitations

Several delimitations should be acknowledged. *First*, the study focused exclusively on fourth-year Japanese-major students at the USSH, VNU-HCM. *Second*, the investigation centered on students' perceptions and experiences rather than objective measures of language proficiency. *Third*, the bilingual instructional model was implemented within a Japanese Reading Comprehension course and may not reflect learning conditions in other language courses or institutions.

In addition, the relatively small sample size and the use of self-reported data may limit the generalizability of the findings. Future research could incorporate larger participant groups, longitudinal designs, and direct assessments of language performance to provide a more comprehensive evaluation of ESL integration in multilingual learning environments.

3. Discussion and findings

3.1. A multimodal and bilingual approach in the classroom

The intervention implemented in this research was designed as a student-centered, multimodal environment that prioritizes genuine understanding over rote translation. In this model, the classroom is divided into small groups, each responsible for presenting a specific reading passage. The presentation phase follows a "non-Vietnamese" guideline for explanation, but participation in English use remains voluntary. When defining new vocabulary or complex Kanji, the presenting group is encouraged to avoid the mother tongue and to experiment with Japanese definitions, Japanese synonyms, visual aids, realia, gestures, and, where possible, English equivalents to convey meaning. The ultimate aim is not to force English into the classroom, but to help students realize that once Vietnamese is set aside, they can flexibly draw on multiple resources (especially their long-acquired L2 English, to interpret and explain the L3). This force-free, multimodal strategy encourages students to engage with Japanese through various cognitive channels, reducing the immediate reliance on Vietnamese.

Following the vocabulary presentation, the rest of the class participates in a "live translation" challenge. This phase is designed to be low-pressure yet intellectually stimulating. Students are invited to translate the reading content into Vietnamese or, as a preferred challenge, into English without the aid of dictionaries. This spontaneous translation process is intended to test their immediate linguistic reflexes. During this time, both the presenting group and the teacher act as facilitators, monitoring the "live" outputs to provide immediate feedback on accuracy and nuance.

The final stage of this model involves a collaborative confirmation of the text's meaning. After the students have attempted their translations, the teacher provides the "standard" or polished version prepared by the students in advance. This ensures that while the process is creative and bilingual, the final learning outcome remains academically precise. By treating English not as a mandatory requirement but as an optional tool for those who wish to experiment, this paper argues that the classroom can become a more dynamic space, where Japanese texts are interprted through a broader, more international perspective, not only through Vietnamese.

3.2. Results and findings

3.2.1. Participant profile and engagement rate

This research was conducted within the fourth-year cohort at the Department of Japanese Studies (USSH, VNU-HCM), where the typical class size ranges from 25 to 32 students across three groups. Out of a total population of approximately 85 students, this paper successfully gathered 51 valid responses, achieving a 60% participation rate. This level of engagement provides a reliable snapshot of the student body's diverse linguistic backgrounds.

The findings revealed a significant "proficiency gap" among the participants. While a majority (72.5%) possess a solid N2 level in Japanese, their English proficiency (L2) is far less uniform. Approximately 45% of respondents are at or below the Pre-intermediate level (IELTS 5.0 equivalent), while only 14% have reached the Upper-intermediate or Advanced stages. This diversity is essential for this paper's analysis, as it highlights that the effectiveness of the classroom intervention is not one-size-fits-all but varies according to individual linguistic capital (Hulstijn, 2015). This proficiency imbalance is particularly relevant to the theoretical assumptions underlying the study. From the perspective of language attrition, the findings suggest that English proficiency may become increasingly vulnerable when learners devote substantial attention to Japanese language development. At the same time, the diversity of English proficiency levels provides an appropriate context for examining whether students'

responses to ESL integration vary according to their existing linguistic resources. The distribution of English proficiency levels is illustrated in **Figure 1**.

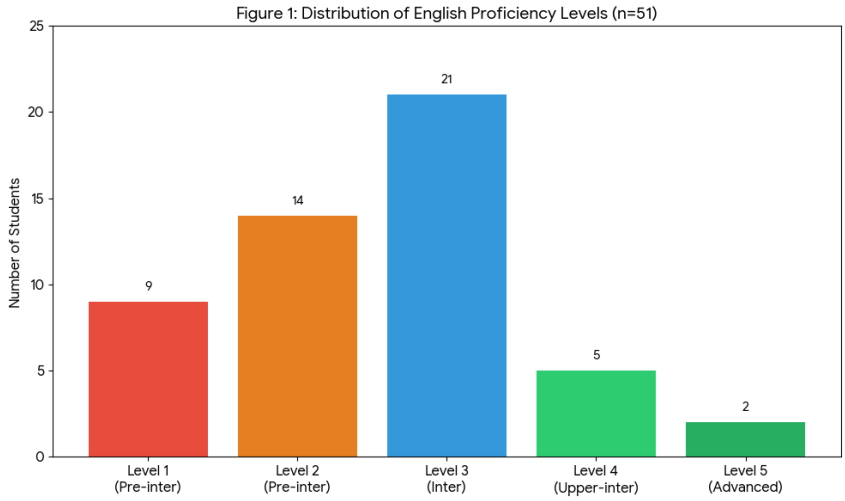


Figure 1. Distribution of English proficiency levels among the surveyed students (n=51)

Source: Compiled by the researcher from survey data.

3.2.2. Quantitative findings

The descriptive statistics indicate a generally high level of awareness regarding English attrition among the participants. The mean score for perceived decline in English proficiency (M=4.18, SD=1.05) reveals that many students recognized a reduction in opportunities to actively use English while concentrating on Japanese language studies. Similarly, the relatively high score for difficulty in switching between Japanese and English (M=4.26, SD=0.89) reflects the linguistic challenges associated with managing multiple foreign languages simultaneously.

Participants also reported generally positive perceptions of the bilingual instructional model. Scores related to vocabulary retention through English equivalents (M = 3.90, SD = 0.98), perceived enjoyment (M=3.63, SD=0.87), and excitement in identifying cross-linguistic connections (M=3.76, SD=0.93) suggest that students viewed ESL integration as a potentially valuable learning strategy. These findings point to a cross-linguistic transfer mechanism consistent with Cummins' *Interdependence Hypothesis*, in which prior linguistic knowledge is reused across languages.

Moreover, the highest mean scores were observed in items related to career readiness and employability. Participants strongly agreed that multilingual competence represented an important professional asset, as reflected in the high ratings for confidence in applying to multinational companies (M=4.37, SD=0.72) and the perceived importance of the "English + 1" model (M = 4.36, SD=1.04). These findings suggest that students viewed ESL integration not only as a learning strategy but also as a means of enhancing their future professional competitiveness. Descriptive statistics for the key variables are summarized in **Table 2**.

Table 2: Descriptive statistics of student perception (n=51)

Survey items (Key variables)	**Mean (M)**	**Std. Deviation (SD)**
A. Language attrition awareness		

1. Perception of English skills "fading"	4.18	1.05
2. Difficulty in switching between JPN and ENG	4.26	0.89
B. Bilingual model effectiveness		
3. Vocabulary retention via equivalents	3.90	0.98
4. Interesting challenge (instead of pressure)	3.63	0.87
5. Excitement in finding linguistic connections	3.76	0.93
C. **Practicality & career**		
6. Confidence in applying to MNCs	4.37	0.72
7. Necessity of "English + 1" model	4.36	1.04

Source: Calculated from the primary survey data (2026) by the author

A one-way ANOVA was conducted to examine whether students' perceptions differed according to English proficiency level. The analysis revealed a statistically significant difference in perceived interest across proficiency groups ($p < .05$). Students with higher levels of English proficiency reported more positive attitudes toward the instructional model, whereas learners in the lowest proficiency group expressed comparatively lower levels of interest and engagement.

This finding supports the *Language Threshold Hypothesis*, which suggests that learners must possess a minimum level of linguistic competence before they can effectively benefit from cross-linguistic transfer. For students with stronger English proficiency, the bilingual instructional model appeared to function as a cognitively stimulating learning strategy that facilitated connections between Japanese and English. In contrast, learners with more limited English proficiency may have experienced difficulties managing the simultaneous activation of two foreign languages.

The findings also align with *Cognitive Load Theory*. While multilingual interaction can promote deeper processing and transfer of knowledge, excessive cognitive demands may hinder learning when students lack sufficient linguistic resources. Consequently, the effectiveness of ESL integration appears to depend not only on the instructional design itself but also on learners' existing proficiency levels. Differences in perceived interest across proficiency levels are illustrated in **Figure 2**.

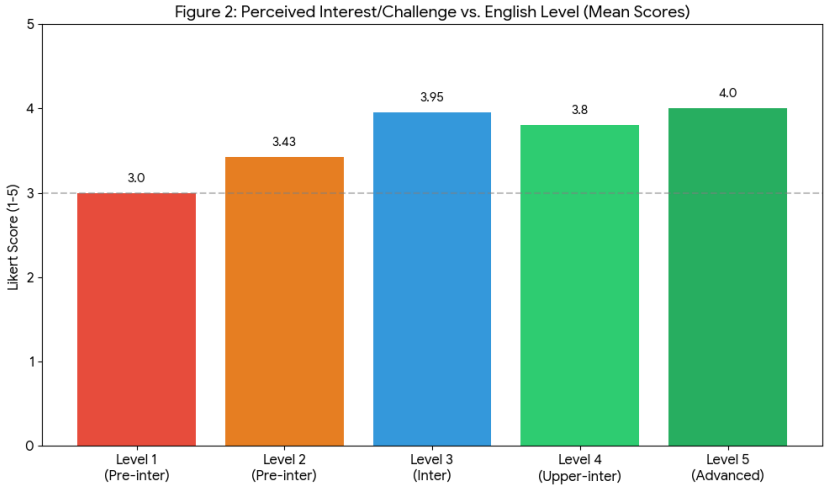


Figure 2. *Mean scores of Perceived Interest versus English proficiency levels*

Source: Extracted from SPSS analysis of survey responses 2026

3.2.3. Qualitative findings

The qualitative responses provided valuable insights into how students experienced the integration of English into Japanese reading instruction. Through thematic analysis, three major themes emerged, which were English as a cognitive bridge, enhanced multilingual awareness, and cognitive burden associated with linguistic limitations. This thematic analysis revealed three recurring themes in students' responses: English as a cognitive bridge, enhanced multilingual awareness and learning motivation, and cognitive burden associated with linguistic limitations.

Theme 1: English as a cognitive bridge for Japanese reading comprehension

A recurring theme across participants' responses was the perception of English as a useful intermediary for understanding Japanese vocabulary and grammatical structures. Many students reported that identifying English equivalents enabled them to process unfamiliar Japanese expressions more efficiently and establish meaningful connections between previously acquired knowledge and new linguistic input.

This finding supports Cummins' *Interdependence Hypothesis*, which proposes that knowledge acquired in one language can facilitate learning in another through shared learning capacity. Rather than functioning as separate linguistic systems, English and Japanese appeared to interact dynamically within students' learning processes. For many participants, English served as a practical scaffold that supported comprehension and promoted deeper engagement with Japanese texts.

Theme 2: Development of multilingual awareness and learning motivation

A second theme concerned the development of multilingual awareness. Participants frequently described feelings of satisfaction when recognizing similarities among languages and successfully drawing connections between Japanese and English. Several students reported that the instructional approach encouraged them to think more critically about language structures and vocabulary relationships rather than relying exclusively on memorization.

These experiences are consistent with recent perspectives on translanguaging and multilingual education, which emphasize the pedagogical value of drawing upon learners' full

linguistic repertoires (Garcia & Wei, 2014; Cenoz & Gorter, 2022). The findings suggest that ESL integration may promote a more flexible and reflective approach to language learning by encouraging students to utilize multiple linguistic resources simultaneously. Furthermore, the ability to engage with more than one foreign language appeared to strengthen students' confidence in their future academic and professional development.

Theme 3: Cognitive burden and linguistic insecurity

Despite the generally positive responses, a third theme highlighted challenges associated with the bilingual instructional model. Participants with lower levels of English proficiency frequently reported feelings of confusion, mental fatigue, and anxiety when attempting to process Japanese through English. For these students, the additional requirement to activate English sometimes increased the complexity of the learning task rather than facilitating comprehension.

This finding can be interpreted through the lens of *Cognitive Load Theory* (Paas & Ayres, 2014), which suggests that excessive demands on working memory may negatively affect learning performance. The results also provide support for the *Language Threshold Hypothesis* (Cummins, 2017), indicating that learners may require a certain level of English proficiency before they can fully benefit from cross-linguistic transfer. Therefore, the benefits of ESL integration may vary across learners, particularly according to their English proficiency levels.

Overall, the qualitative findings complement the quantitative results by revealing both the opportunities and challenges associated with ESL integration in Japanese reading instruction. While many students perceived English as valuable cognitive resources that enhanced multilingual awareness and reading comprehension, others experienced difficulties resulting from limited English proficiency. These findings suggest that the implementation of bilingual instructional practices should remain flexible and responsive to learners' varying linguistic backgrounds. In this sense, the model appears more suitable for students who already possess a relatively stable English foundation, which may help reactivate and sustain their L2 while supporting Japanese reading comprehension.

4. Conclusions and pedagogical implications

The present study investigated the integration of English as a supporting language in Japanese reading instruction among fourth-year Japanese-major students at the USSH, VNU-HCM. Drawing upon theories of language attrition, cross-linguistic transfer, translanguaging, and cognitive load, the study explored students' perceptions of a bilingual instructional model designed to maintain English engagement while supporting Japanese reading comprehension.

The findings indicate that participants generally perceived the integration of English positively. Many students reported that English functioned as a useful cognitive bridge, enabling them to establish meaningful connections between previously acquired linguistic knowledge and new Japanese vocabulary and structures. The results further suggest that ESL integration may contribute to reducing the effects of language attrition by creating opportunities for students to continue using English within their academic learning environment.

At the same time, this study revealed important differences among learners. Students with stronger English proficiency demonstrated greater enthusiasm and adaptability toward the bilingual instructional approach, whereas those with more limited English backgrounds were more likely to experience confusion, frustration, and increased cognitive burden. These

findings support both the *Language Threshold Hypothesis* and *Cognitive Load Theory*, suggesting that the effectiveness of multilingual instructional practices is influenced by learners' existing linguistic resources.

Overall, in this study, students reported that integrating English into Japanese reading instruction may provide a practical way to support multilingual language development among Japanese major students. Rather than treating English and Japanese as competing languages, the study supports the view that they can function as complementary resources that mutually reinforce language learning when implemented appropriately.

Several pedagogical implications can be drawn from the findings of this study. *First*, ESL integration should be implemented gradually and adapted to students' English proficiency levels. While higher-proficiency learners may benefit from more extensive English-mediated activities, lower-proficiency students may require additional scaffolding to prevent excessive cognitive load. *Second*, instructors should adopt translanguaging practices that encourage flexible movement between Japanese, English, and Vietnamese rather than enforcing rigid language separation. Such practices may help learners utilize their full linguistic repertoire and facilitate deeper comprehension of reading materials. *Third*, vocabulary instruction can be enhanced through the systematic incorporation of English-Japanese lexical mapping activities. Encouraging students to identify semantic relationships among languages may strengthen vocabulary retention and promote cross-linguistic awareness. *Fourth*, multilingual reading tasks should emphasize meaning-making rather than direct translation. Activities such as collaborative discussion, vocabulary comparison, and contextual interpretation may help students engage more actively with texts while maintaining meaningful use of English. *Finally*, universities may consider incorporating multilingual instructional approaches into language programs as a means of supporting graduates' employability. In an increasingly globalized labor market, maintaining proficiency in both English and Japanese may provide students with a significant professional advantage.

Several limitations should be acknowledged. *First*, this study was conducted with a relatively small sample drawn from a single institution, which may limit the generalizability of the findings. *Second*, this investigation relied primarily on self-reported perceptions rather than objective measures of language proficiency. *Third*, this intervention was implemented within a specific Japanese reading course and over a limited period of time.

Future studies may expand the sample size, include participants from multiple institutions, and employ longitudinal research designs to examine the long-term effects of ESL integration on both English maintenance and Japanese language development. Furthermore, future research could incorporate standardized language assessments to provide more objective evidence regarding the effectiveness of multilingual instructional approaches.

Form the teacher's point of view, the bilingual model also made classroom interaction more lively, as students had more options for expressing partial understanding before switching back to Japanese.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Cenoz, J., & Gorter, D. (2022). Pedagogical translanguaging and its application to language classes. *RELC Journal: A Journal of Language Teaching and Research, 53*(2), 342-354. https://doi.org/10.1177/00336882221082751

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.

Cummins, J. (2017). *Bilingual education and bilingualism* (6th ed.). Multilingual Matters.

Cummins, J. (2021). *Rethinking the education of multilingual learners*. Multilingual Matters.

Garcia, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Multilingual Matters.

Hulstijn, J. H. (2015). *Language proficiency in native and non-native speakers: Theory and research*. John Benjamins.

Kopke, B., & Schmid, M. S. (2004). Language attrition: The next phase. In M. S. Schmid, B. Kopke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* 1-43. John Benjamins.

Nguyen, X. N. C. M., & Nguyen, V. H. (2019). Language education policy in Vietnam. In A. Kirkpatrick & A. J. Liddicoat (Eds.), *The Routledge international handbook of language education policy in Asia,* 185-201. Routledge. https://doi.org/10.4324/9781315666235-13

Paas, F., & Ayres, P. (2014). Cognitive load theory: A broader view on the role of memory in learning and education. *Educational Psychology Review*, 26(2), 191-195.

Schmid, M. S., & Kopke, B. (2017). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, *7*(6), 637-667. https://doi.org/10.1075/lab.17058.sch

Schmid, M. S., & Kopke, B. (Eds.). (2019). *The Oxford handbook of language attrition. Oxford University Press.* https://doi.org/10.1093/oxfordhb/9780198793595.001.0001.

Taber, K. S. (2017). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Wei, L. (2017). Translanguaging as a practical theory of language. *Applied Linguistics*, 39(1), 9-30. https://doi.org/10.1093/applin/amx039

#85 FACTORS ASSOCIATED WITH THE SCIENTIFIC RESEARCH COMPETENCE OF HUFLIT ENGLISH LANGUAGE STUDENTS IN PROJECT 1 AND PROJECT 2

Quoc Bao Mai[1]

Abstract

This study examines factors associated with the self-assessed scientific research competence (SRC) of English Language students at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT) through Project 1 and Project 2. The study uses a convergent mixed-methods design, combining a quantitative survey of 500 students who had completed both courses with semi-structured interviews with 19 students. Quantitative data were processed with SPSS 26.0 and SmartPLS 4.0 using PLS-SEM, while qualitative data were analyzed thematically. The results show that the measurement model achieved satisfactory reliability, convergent validity, and discriminant validity (HTMT < 0.85). The structural model explained 52.7% of the variance in self-assessed SRC, 37.5% of the variance in overall course effectiveness (OE), and 46.0% of the variance in satisfaction (SAT). Of the 12 hypotheses, 11 were supported. Learning engagement (PE) was the strongest predictor of SRC ($\beta = 0.279$, $p < .001$), followed by assessment and feedback standardization (AS), course design quality and continuity (AL_CW), lecturer/system support (IS), and AI guidance linked with academic integrity (AI). Methodological and academic language barriers (MLB) showed a positive but statistically non-significant relationship with SRC ($\beta = 0.010$, $t = 0.342$, $p = .733$), contrary to the hypothesized negative direction. The interview data suggest that this result may reflect students' awareness of research difficulties and their use of coping strategies (such as AI or peer reliance), rather than only the actual obstructive level of those difficulties. Students still reported difficulties with literature review, SPSS-based data analysis, discussion writing, teamwork, and responsible AI use. The findings provide evidence for improving course design, standardizing rubrics, strengthening formative feedback, and developing clearer guidelines for AI use in academic work.

Keywords: self-assessed scientific research competence, English Language students, research methodology courses, PLS-SEM, assessment and feedback standardization, AI guidance, academic integrity

1. Introduction

1.1. Rationale

In the context of modern higher education, scientific research competence is not only an academic requirement but also an essential skill that helps students develop critical thinking, problem-solving ability, and lifelong learning (Ciraso-Cali et al., 2022; Prosekov et al., 2020). For English Language students, doing scientific research has a more specific and complex nature. They face a "double burden": they must understand research methodology and also

[1] *Corresponding author, Undergraduate Student, Department of English Language, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT), Ho Chi Minh City, Vietnam, Email: 23DH710401@st.huflit.edu.vn; ORCID: 0009-0002-3382-0809*

overcome barriers in academic language, data analysis, and culture (Ciraso-Cali et al., 2022; Le et al., 2023).

Recognizing this importance, Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT) has included two core courses in the English Language program: Project 1 (Data Collection Skills) and Project 2 (Introduction to Scientific Research Methodology). The design of this course sequence shows the university's effort to move from content transmission to outcome-based education, creating constructive alignment in teaching (Biggs, 1996), together with a project-based learning model.

However, in practice, the implementation of research courses still has many challenges. Students often have difficulty identifying research problems, designing instruments, and especially meeting the strict standards of academic English. Moreover, the continuity between the two courses, teaching quality, supporting learning materials, and consistency in assessment through clear criteria are factors that directly are related to the success of developing students' research competence (Brookhart, 2018; Dawson, 2015; Reddy & Andrade, 2009).

Previous studies on students' research competence have mainly focused on describing competence or motivation to participate in research. They have not fully explained the combined roles of course design, lecturer support, assessment rubrics, AI/academic integrity guidance, and language/methodological barriers in the context of two connected Project courses. To address this gap, the topic "Factors Associated with the Scientific Research Competence of HUFLIT English Language Students in Project 1 and Project 2" was conducted.

1.2. Research Objectives

This study aims to identify and analyze the factors associated with the scientific research competence of English Language students at HUFLIT through Project 1 and Project 2. Specifically, the study examines the relationships of learning engagement, lecturer and system support, assessment and feedback standardization, AI use and academic integrity guidance, methodological/language barriers, and course design quality and continuity on students' SRC. Based on quantitative and qualitative findings, the study proposes implications for course improvement and directions for building an SRC framework suitable for English Language students.

1.3. Research Questions

To achieve these objectives, the study focuses on answering three questions:

- What factors are associated with the self-assessed scientific research competence of HUFLIT English Language students in Project 1 and Project 2?
- What is the strength of the relationships between learning engagement, lecturer/system support, assessment and feedback standardization, AI guidance, methodological/language barriers, course design quality, and students' SRC?
- How is students' SRC associated with overall course effectiveness and student satisfaction?

1.4. 1.4. Significance of the Study

Theoretical significance:

This study enriches the academic literature on the scientific research competence of university students, especially in the context of English as a Foreign Language (EFL). The use of a multi-dimensional assessment model (survey and interview) provides a strong mixed-methods approach through triangulation in educational research (Creswell & Creswell, 2009).

Practical significance:

- For students: The study helps students clearly identify difficulties in doing research, including both methodology and academic English, so that they can develop more effective self-regulated learning strategies (Panadero & Broadbent, 2018).
- For lecturers and the Faculty of Foreign Languages (HUFLIT): The study provides empirical evidence for reviewing and adjusting syllabi, standardizing assessment criteria (Brookhart & Chen, 2014; Dawson, 2015), and strengthening the continuity between Project 1 and Project 2.
- For administrators and curriculum developers: The study provides empirical evidence for the university to review and improve the training program, especially the continuity between Project 1 and Project 2, supporting learning materials, assessment rubrics, feedback mechanisms, and guidance on AI use in academic work. On this basis, the program can be adjusted to better match the learning outcomes of the English Language major and the need to develop students' scientific research competence.

2. Literature Review

2.1. Root Theories

To explain the development of SRC and the factors associated with it, this study is based on the following main theories:

2.1.1. Theory of Planned Behavior (TPB)

The Theory of Planned Behavior, proposed by Ajzen (1991), is one of the classic theories used to predict individual intention and behavior. According to TPB, students' research behavior and SRC are affected by three factors: (1) attitude toward the behavior, (2) subjective norms, such as encouragement and support from lecturers and the university, and (3) perceived behavioral control, or students' confidence in overcoming barriers. This theory supports the argument that institutional scaffolding and language/methodological barriers are directly related to students' ability to conduct research.

2.1.2. Constructive Alignment and Outcome-Based Education (OBE)

Constructive Alignment, developed by Biggs (1996), is an important guide for higher education curriculum design. Biggs argues that learning is most effective when there is close alignment among learning objectives or outcomes, teaching methods, and assessment methods. For Project 1 and Project 2, this theory explains why students' SRC may depend strongly on the continuity of course design and the transparency of assessment criteria, or rubrics.

2.1.3. Self-Regulated Learning (SRL)

According to Panadero and Broadbent (2018), self-regulated learning is the process in which learners guide, monitor, and adjust their behavior and thinking to achieve academic goals. In project-based learning environments such as Project 1 and Project 2, active learning engagement and self-assessment are internal drivers that help students turn passive knowledge into real research competence.

2.2. Empirical Literature

2.2.1. Scientific Research Competence and the "Double Burden" in an EFL Context

SRC is a multi-dimensional system that includes the ability to identify problems, design methods, collect data, analyze data, and present findings (Ciraso-Cali et al., 2022). Compared with students in natural science fields, English Language students face a "double burden". First,

they face methodological barriers when learning data analysis concepts (Ciraso-Cali et al., 2022; Davidson & Palermo, 2015). Second, they face academic language barriers. The gap between everyday communication and academic language creates major difficulties in reading academic sources and writing standard reports (Le et al., 2023).

2.2.2. The Role of Assessment for Learning (AfL), Feedback, and Rubrics

Assessment for Learning (AfL) shifts the focus from merely grading students to using assessment as a tool to enhance their learning process (Black & Wiliam, 1998). In complex tasks like scientific research, effective formative feedback provides students with the necessary scaffolding to identify their weaknesses and improve their work iteratively (Hattie & Timperley, 2007). Furthermore, the use of clear and well-designed rubrics is essential in higher education. Rubrics not only standardize assessment criteria among different lecturers but also help students understand academic expectations, thereby developing their evaluative judgement and self-regulation skills (Brookhart, 2018; Dawson, 2015; Reddy & Andrade, 2009).

2.2.3. Artificial Intelligence (AI) and Academic Integrity in Research

The growth of digital tools, open science, and AI is reshaping research practice. Requirements for transparency, open data, and research integrity are becoming more important (Le et al., 2025). Appropriate AI use can support students in data processing and academic writing. However, without clear guidance on academic integrity, AI may replace learners' critical thinking. Therefore, integrating AI guidance into the Project courses is an urgent need.

2.3. Hypothesis Development

Drawing on the theoretical and empirical review above, this study constructs a structural model centering on SRC as the outcome variable. It is predicted by six input factors and then linked to overall course effectiveness and student satisfaction. The 12 hypotheses below are organized in two groups.

Group 1: Factors Affecting Scientific Research Competence (SRC)

Self-Regulated Learning theory treats active learner agency as the primary catalyst for competence development (Panadero & Broadbent, 2018). Healey et al. (2014) further argue that scientific research is inherently participatory: students cannot acquire research skills through passive reception of lecture content alone; rather, they must engage with materials, interact with lecturers and peers, and monitor their own progress over time. Such engagement transforms declarative knowledge about research methods into procedural competence. In the context of Project 1 and Project 2, where students are expected to produce extended research outputs, high levels of engagement - including autonomous resource-seeking, deadline management, and iterative revision - should be associated with a stronger command of the research process.

H1 (+): Active learning engagement (PE) is positively associated with students' SRC.

Vygotsky's concept of the zone of proximal development implies that novice researchers benefit most when guided by knowledgeable others who provide timely scaffolding (Ciraso-Cali et al., 2022; Le et al., 2023). In the HUFLIT context, this scaffolding takes two distinct forms: institutional resources (e.g., Moodle repositories, research templates, handbooks) and supervisory guidance (e.g., formative feedback on draft chapters, consultations on methodology). When either form is insufficient or delayed, students may lose direction at

critical stages such as instrument design or data interpretation, ultimately undermining SRC development.

H2 (+): System and lecturer support (IS) is positively associated with SRC.

The literature on assessment for learning establishes that well-designed rubrics serve a dual pedagogical function: they communicate academic expectations transparently and they provide a diagnostic lens through which students can evaluate - and subsequently improve - their own work (Black & Wiliam, 1998; Brookhart, 2018; Dawson, 2015; Hattie & Timperley, 2007). When rubric criteria are applied consistently across lecturers, students develop evaluative judgement - the capacity to discern quality and to close the gap between current and target performance. By contrast, ambiguous or inconsistent grading erodes trust in the assessment process and can discourage sustained effort.

H3 (+): Assessment and feedback standardization (AS) is positively associated with SRC.

The emergence of generative AI tools (e.g., ChatGPT, Gemini) has introduced a new dimension to research pedagogy. While AI can accelerate literature synthesis, grammar correction, and data exploration, it simultaneously raises concerns about intellectual autonomy and academic honesty (Le et al., 2025). Whether AI supports or undermines SRC depends largely on how students are instructed to use it. Clear guidelines that delineate acceptable assistance (e.g., brainstorming prompts, language polishing) from academic misconduct (e.g., submitting AI-generated analysis without understanding) enable students to leverage digital tools responsibly, thereby enhancing rather than substituting their research competence.

H4 (+): AI use and academic integrity guidance (AI) is positively associated with SRC.

English Language students confront a "double burden": limited statistical training and the formidable task of producing academic discourse that meets international publication standards (Ciraso-Cali et al., 2022; Le et al., 2023). Within the Theory of Planned Behavior, these barriers reduce perceived behavioral control -- i.e., the learner's belief that they can successfully execute the research task -- which in turn dampens motivation and may constrain final competence. Accordingly, students who perceive greater methodological and linguistic obstacles should, on average, report lower SRC.

H5 (-): Methodological and academic language barriers (MLB) are negatively associated with students' SRC.

Constructive Alignment (Biggs, 1996) stipulates that effective learning requires tight coherence among intended outcomes, instructional activities, and assessment tasks. When applied to a two-course sequence, this principle further demands vertical alignment: the competencies developed in Project 1 should serve as explicit prerequisites for the more demanding tasks of Project 2. A well-articulated progression -- coupled with manageable workload distribution -- enables students to consolidate foundational skills before tackling full-scale research, thereby strengthening SRC.

H6 (+): Course design quality and continuity (AL_CW) is positively associated with SRC.

Group 2: Effects on Satisfaction (SAT) and Overall Effectiveness (OE)

From an Outcome-Based Education perspective, course success is ultimately measured by the competencies that learners acquire. Students who perceive genuine growth in their research capabilities -- from formulating research questions through to interpreting findings and writing academic reports -- are more likely to judge the course as effective and to express satisfaction with their academic investment.

H7 (+): Students' SRC is positively associated with overall course effectiveness (OE).

H8 (+): Students' SRC is positively associated with satisfaction with the course (SAT).

Learner satisfaction is not determined solely by acquired competence; it is also shaped by the quality of the learning experience itself. A curriculum that is logically sequenced, offers appropriate workload, and maintains clear expectations contributes positively to both satisfaction and perceived effectiveness (Reddy & Andrade, 2009). Lecturer and system support create a responsive learning environment, while transparent assessment gives students a sense of procedural fairness and predictability. These institutional factors therefore merit separate paths to the outcome variables:

H9 (+): Course design quality and continuity (AL_CW) is positively associated with student satisfaction (SAT).

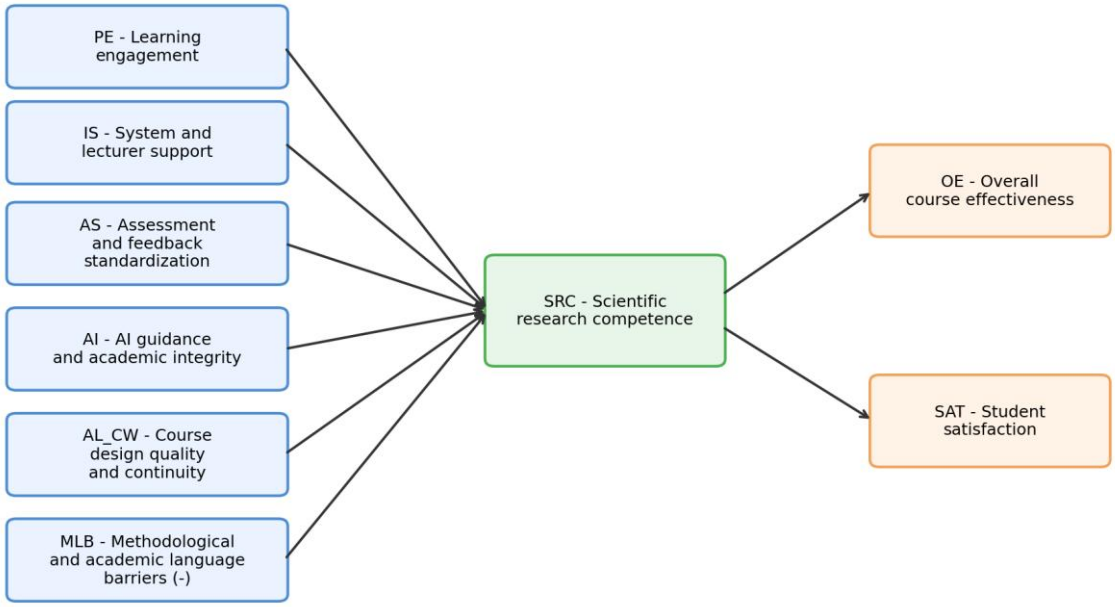
H10 (+): System/lecturer support (IS) is positively associated with student satisfaction (SAT).

H11 (+): Course design quality and continuity (AL_CW) is positively associated with overall course effectiveness (OE).

H12 (+): Assessment and feedback standardization (AS) is positively associated with student satisfaction (SAT).

Figure 1

Research model of the proposed effects



3. Methodology

3.1. Research Model and Rationale for Factor Selection

A growing body of evidence suggests that students' SRC is a product not only of individual aptitude but also of the broader academic ecosystem -- including curriculum design, supervisory quality, resource availability, and assessment culture. In the Vietnamese higher-education context, Tinh et al. (2020) identify research awareness, participation motivation, self-assessment capability, and the institutional support environment as key conditions shaping

scientific research capacity. Complementary work on Vietnamese undergraduates further highlights the interplay of university-level encouragement, learning conditions, personal competence, and intrinsic motivation (Bui & Tran, 2022; Son & Mai, 2021; Vo, 2023).

Building on these foundations, the present study constructs an integrated model situating students' self-assessed SRC at the center of a network of six predictor groups -- active learning engagement (PE), lecturer and system support (IS), assessment and feedback standardization (AS), AI use and academic integrity guidance (AI), methodological and academic language barriers (MLB), and course design quality and continuity (AL_CW) - together with two outcome variables: overall course effectiveness (OE) and student satisfaction (SAT). Direct paths from AL_CW, IS, and AS to SAT and OE are also specified. This design permits a simultaneous evaluation of the individual, pedagogical, and institutional factors that may support or constrain SRC development within the two-course Project sequence.

3.2. Research Design

The study employs a convergent mixed-methods design (Creswell & Creswell, 2009). The quantitative strand relies on a structured questionnaire to test the theoretical model and examine hypothesized relationships among the nine latent variables. Concurrently, the qualitative strand draws on semi-structured interviews to elicit richer accounts of students' lived experiences, practical barriers, and perceived support needs across Project 1 and Project 2. Integrating both data sources serves two purposes: it mitigates the inherent limitation of exclusive reliance on self-report measures, and it produces a more complete understanding of the mechanisms through which SRC is developed.

3.3. Research Procedure

Five sequential stages structured the investigation. In the first stage, the researcher reviewed theoretical frameworks (TPB, SRL, Constructive Alignment) and empirical studies to identify the research gap, formulate the structural model, and derive the twelve hypotheses. The second stage involved instrument development: both the quantitative questionnaire and the semi-structured interview guide were drafted, subjected to a pilot survey, and refined to ensure contextual appropriateness for English Language students at HUFLIT. During the third stage, data were collected through a large-scale survey and in-depth interviews conducted simultaneously. In the fourth stage, quantitative data were analyzed in SPSS 26.0 (descriptive statistics, sample profiling) and SmartPLS 4.0 (PLS-SEM, bootstrapping, CTA-PLS), while qualitative transcripts were processed through open, axial, and selective coding in MAXQDA. The fifth and final stage synthesized both datasets, comparing quantitative path results with interview themes to generate conclusions, course-improvement implications, and recommendations for an SRC framework.

3.4. Research Sample and Sampling Method

The quantitative sample comprised 500 English Language students at HUFLIT who had completed both Project 1 and Project 2 - a criterion ensuring that all respondents possessed sufficient learning experience to evaluate course design, support structures, assessment practices, and their own SRC development. Sampling followed a two-stage procedure: convenience sampling was used to identify eligible cohorts, followed by random selection within Project classes to enhance representativeness and meet the requirements of PLS-SEM analysis. For the qualitative component, 19 students were purposively selected to maximize variation in learning outcomes, course experiences, and major orientations.

Regarding demographic composition, female students accounted for 55.2% and male students for 44.8%. Major orientations were distributed as follows: Office Administration (18.0%), Translation and Interpretation (17.6%), English for Logistics (17.6%), Business English (17.2%), English Language Teaching (15.4%), and English-Chinese Bilingual Studies (14.2%). All participants had taken both Project courses, confirming the suitability of the sample for the proposed model.

Table 1

Demographic characteristics of the survey sample

Characteristic	Category	Frequency (n)	Percentage (%)
Gender (GI1)	Male	224	44.8
	Female	276	55.2
Major orientation (GI3)	Translation and Interpretation	88	17.6
	English Language Teaching	77	15.4
	Business English	86	17.2
	English-Chinese Bilingual Studies	71	14.2
	English for Logistics	88	17.6
	Office Administration	90	18.0
Self-assessment of learning outcomes (GI6)	Very good	83	16.6
	Good	103	20.6
	Average	114	22.8
	Somewhat difficult	111	22.2
	Great difficulty	89	17.8

3.5. Scale and Questionnaire Development

The quantitative questionnaire was designed based on the main theories and previous studies, and then adjusted to fit the context of English Language students at HUFLIT. The questionnaire included a demographic information section and a latent-variable scale section using a 5-point Likert scale. The variables in the model included learning engagement (PE), lecturer and system support (IS), assessment and feedback standardization (AS), AI use and academic integrity guidance (AI), methodological and academic language barriers (MLB), course design quality and continuity (AL_CW), self-assessed scientific research competence (SRC), overall course effectiveness (OE), and student satisfaction (SAT).

Self-assessed SRC was measured through core components that reflect the process of conducting an academic study in the English Language context. These components included identifying a research problem and research questions, reviewing the literature, designing the research method and instrument, collecting data, analyzing and interpreting data, writing a research report in academic English, presenting research findings, and cooperating during the

project. This variable reflects how students evaluate their own competence after completing Project 1 and Project 2. Meanwhile, AI use and academic integrity guidance (AI) was treated as a separate independent variable to avoid conceptual overlap with SRC.

For qualitative data, the semi-structured interview guide was developed around key themes, including learning experiences, methodological/language barriers, the role of supervisors, the usefulness of rubrics, AI use in academic work, and suggestions for course improvement.

3.6. Data Collection Method

Quantitative data were collected through Google Forms and distributed directly in classes as well as through internal learning channels of English Language students. The survey was conducted after students had completed both Project courses to ensure that participants had enough experience to evaluate them. Qualitative data were collected through face-to-face or online interviews, with an average duration of 20 to 30 minutes for each student. The interviews were recorded with participants' consent and transcribed into text for analysis.

3.7. Data Analysis Method

Quantitative data were first processed using SPSS 26.0 for descriptive statistics of the survey sample. Then, SmartPLS 4.0 was used to analyze the partial least squares structural equation model (PLS-SEM). The measurement model was assessed through internal consistency reliability, convergent validity, and discriminant validity. Reliability was examined using Cronbach's alpha, rho_A, and Composite Reliability, with the accepted threshold above 0.70. Convergent validity was assessed using factor loading and Average Variance Extracted (AVE). In principle, factor loading should be about 0.70 or higher. Indicators close to 0.70 may be retained if Composite Reliability and AVE meet the required level. AVE should be 0.50 or higher.

Discriminant validity of the measurement model was assessed using the Fornell-Larcker and HTMT criteria. According to the Fornell-Larcker criterion, the square root of the AVE of each latent variable must be higher than its correlations with the other latent variables. In addition, the HTMT ratio (Heterotrait-Monotrait Ratio), following Henseler et al. (2014), was used to test the level of distinction among concepts in the model. An HTMT value below 0.85 is considered acceptable at a strict level, while the 0.90 threshold may be accepted for closely related concepts. Using both criteria helps ensure that the variables in the measurement model are conceptually different before the structural model is analyzed.

The structural model was assessed using the collinearity coefficient VIF, the coefficient of determination R-squared, the effect size f-squared, the predictive relevance Q-squared, and hypothesis testing with bootstrapping using 5,000 subsamples. A hypothesis was supported when the path coefficient was statistically significant with p-value < 0.05 and t-statistics > 1.96. Because the study used cross-sectional and self-reported data, the structural paths were interpreted as predictive relationships rather than definitive causal effects. Qualitative data were analyzed through thematic analysis with the support of MAXQDA. This included open coding, axial coding, and selective coding to identify themes corresponding to the variables in the model.

3.8. Research Ethics

The study followed ethical principles in educational research. All participants were informed about the purpose and scope of the study and their right to withdraw at any time without any effect on their learning results. The collected data were used only for academic

purposes. Personal information was anonymized and coded. Citation, data processing, and result reporting were conducted honestly and followed the rules of academic integrity and intellectual property.

4. Results

4.1. Self-assessment of learning outcomes

Table 2 presents students' self-assessment of their learning outcomes after taking Project 1 and Project 2. The results show that the group rating themselves as Average had the highest proportion at 22.8%, followed by the Somewhat difficult group at 22.2%. Meanwhile, 20.6% of students rated their outcomes as Good, 16.6% as Very good, and 17.8% stated that they experienced Great difficulty. This distribution shows that the two Project courses set considerable academic demands for students, especially because they had to combine research skills, data processing, and academic English competence.

Table 2

Self-assessment of learning outcomes

Self-assessment of learning outcomes	Frequency (n)	Percentage (%)
Very good	83	16.6
Good	103	20.6
Average	114	22.8
Somewhat difficult	111	22.2
Great difficulty	89	17.8
Total	500	100.0

4.2. Measurement Model Evaluation

The measurement model was assessed to ensure that the observed variables properly reflected the concepts being measured, or the latent variables. The assessment included reliability, convergent validity, and discriminant validity.

4.2.1. Construct Reliability and Convergent Validity

According to Hair et al. (2019), scale reliability is assessed through Cronbach's Alpha, rho_A, and Composite Reliability (CR), with an accepted threshold of > 0.70. Convergent validity is assessed through factor loading and Average Variance Extracted (AVE), in which AVE should be 0.50 or higher.

Table 3

Results of construct reliability and convergent validity assessment

Latent variable	Cronbach's alpha	rho_a	rho_c	Average Variance Extracted (AVE)	Factor Loading
AI	0.859	0.863	0.905	0.704	.814 - .874
AL_CW	0.943	0.946	0.951	0.663	.758 - .877
AS	0.926	0.928	0.940	0.692	.791 - .881

IS	0.894	0.900	0.922	0.703	.784 - .864
MLB	0.959	0.967	0.964	0.689	.765 - .867
OE	0.901	0.905	0.927	0.718	.808 - .883
PE	0.913	0.915	0.932	0.697	.784 - .865
SAT	0.897	0.897	0.924	0.708	.820 - .865
SRC	0.978	0.978	0.979	0.652	.753 - .876

Based on Table 3, all scales achieved excellent reliability and convergent validity. Cronbach's alpha, rho_A, and rho_C values of all variables were well above the 0.70 threshold, indicating strong internal consistency among the items. Furthermore, the AVE values of all variables exceeded the 0.50 benchmark, ranging from 0.652 to 0.718. The factor loadings presented in the table demonstrate that the observed items effectively reflected their corresponding latent concepts. Therefore, the measurement model robustly met the requirements for further testing of discriminant validity and structural relationships.

4.2.2. Discriminant Validity

Discriminant validity was tested to ensure that the latent variables in the model were conceptually distinct from one another. The results based on the Fornell-Larcker and HTMT criteria are presented in Table 4 and Table 5.

Table 4

Discriminant validity results based on the Fornell-Larcker criterion

	AI	**AL_CW**	**AS**	**IS**	**MLB**	**OE**	**PE**	**SAT**	**SRC**
AI	0.839								
AL_CW	0.501	0.814							
AS	0.494	0.508	0.832						
IS	0.448	0.448	0.480	0.838					
MLB	0.203	0.129	0.147	0.118	0.830				
OE	0.480	0.479	0.537	0.530	0.098	0.847			
PE	0.521	0.527	0.562	0.601	0.153	0.564	0.835		
SAT	0.478	0.479	0.552	0.549	0.108	0.535	0.616	0.842	
SRC	0.499	0.546	0.587	0.537	0.143	0.581	0.632	0.571	0.808

Table 5

Discriminant validity results based on the HTMT ratio (Heterotrait-Monotrait Ratio)

	AI	**AL_CW**	**AS**	**IS**	**MLB**	**OE**	**PE**	**SAT**	**SRC**
AI									
AL_CW	0.554								
AS	0.554	0.542							

IS	0.509	0.486	0.523					
MLB	0.223	0.133	0.153	0.125				
OE	0.544	0.517	0.583	0.587	0.102			
PE	0.587	0.567	0.609	0.664	0.162	0.620		
SAT	0.542	0.519	0.605	0.609	0.116	0.593	0.680	
SRC	0.541	0.567	0.615	0.570	0.142	0.616	0.667	0.608

The discriminant validity results demonstrate that the measurement model successfully met the requirements of both the Fornell-Larcker and HTMT criteria. Specifically, under the Fornell-Larcker criterion, the square root of the AVE of each latent variable (the diagonal values in Table 4) was strictly higher than its correlations with all other variables. More importantly, all HTMT values in Table 5 were well below the stringent 0.85 threshold, with the highest recorded value being just 0.680 (between PE and SAT). This indicates excellent discriminant validity, ensuring that the constructs within the model are empirically distinct and highly suitable for structural equation modeling.

4.2.3. Confirmatory Tetrad Analysis (CTA-PLS)

Traditionally, the decision to set scales as Reflective or Formative is often based only on theoretical arguments. However, to avoid measurement model misspecification, this study conducted Confirmatory Tetrad Analysis (CTA-PLS) to use empirical data to confirm whether the scales used in the study were reflective or formative (Hair et al., 2019).

CTA-PLS tests the statistical significance of tetrads, which are differences between products of covariances of pairs of observed variables. The evaluation criteria are based on the p-value and confidence interval from the bootstrapping procedure. If p-value > 0.05 or the 95% confidence interval includes 0, the study does not have enough evidence to reject the assumption that the scale is reflective. In contrast, if p-value < 0.05 and the confidence interval does not include 0, the scale should be reconsidered as formative.

The SmartPLS results show that, for latent variables with four or more observed variables in the model, the tetrads were generally not statistically significant and the adjusted confidence intervals (CI adjusted) all included 0. This result shows that the empirical data did not reject the reflective specification of the scales in the model.

This result supports the use of reflective scales in the study. It is consistent with the original theoretical basis from the Theory of Planned Behavior (TPB), Self-Regulated Learning (SRL), and the way the model concepts were measured.

4.3. Structural Model Evaluation and Hypothesis Testing

After the measurement model was confirmed, the structural model was analyzed to test the research hypotheses through the bootstrapping algorithm with 5,000 subsamples.

4.3.1. Collinearity Issues

Before testing the effects, the inner VIF was examined to ensure that there was no collinearity problem among the independent variables. Table 5 presents the VIF results.

Table 6

Collinearity test results (VIF)

Independent variable	OE	SAT	SRC
AI			1.638
AL_CW	1.426	1.586	1.641
AS		1.716	1.718
IS		1.532	1.697
MLB			1.048
PE			2.048
SRC	1.426	1.910	

As shown in Table 6, all inner VIF values across the structural model ranged from 1.048 to 2.048. These values are consistently well below the conservative threshold of 3.0 (and far below the maximum acceptable limit of 5.0 set by Hair et al., 2019). This confirms that multi-collinearity was not a problem and did not bias the estimation of the path coefficients.

4.3.2. Explanatory Power (R^2) and Effect Size (f^2)

- R^2 coefficient: This shows the level of variance in the dependent variables (OE, SAT, SRC) explained by the independent variables.
- f^2 coefficient: This evaluates the contribution level, large, medium, or small, of each independent variable to the dependent variable.

Table 7

Explanatory power (R^2) of the model

Dependent variable	R^2	Adjusted R^2
OE	0.375	0.372
SAT	0.460	0.455
SRC	0.527	0.521

Table 8

Effect size (f^2) of the independent variables

Independent variable	OE	SAT	SRC
AI			0.010
AL_CW	0.059	0.017	0.039
AS		0.058	0.065
IS		0.082	0.024
MLB			0.000
PE			0.081

SRC	0.233	0.051

Explanatory power (R-squared): The predictor variables in the model explained 52.7% (R-squared = 0.527) of the variance in Scientific Research Competence (SRC). In educational and behavioral research, this represents a highly satisfactory and realistic level of explanatory power, indicating that the proposed framework comprehensively captures the core factors associated with SRC development. Concurrently, the structural model successfully explained 37.5% of the variance in overall course effectiveness (OE) and 46.0% of the variance in student satisfaction (SAT).

- Effect size (f^2):

  - Among the factors predicting SRC, learning engagement (PE) was the largest contributor with an effect size of $f^2 = 0.081$, representing a small-to-medium meaningful effect.

  - SRC exhibited a medium effect on overall course effectiveness (OE), with $f^2 = 0.233$.

  - Notably, methodological and academic language barriers (MLB) had no practical effect on SRC ($f^2 = 0.000$), which directly aligns with the non-significant outcome observed for hypothesis H5.

4.3.3. Hypothesis Testing

The strength of association between variables was assessed through the path coefficient (beta), t-statistics, and p-value obtained from the bootstrapping process. A hypothesis was accepted when p-value < 0.05 and t-statistics > 1.96.

Table 9

Hypothesis testing results (Bootstrapping)

Hypothesis	Path	Original sample (O)	Mean (M)	Standard deviation (STDEV)	*t* statistics	*p* values	Conclusion
H1	PE -> SRC	0.279	0.281	0.042	6.627	<.001	Accepted
H2	IS -> SRC	0.140	0.140	0.041	3.442	.001	Accepted
H3	AS -> SRC	0.229	0.228	0.039	5.856	<.001	Accepted
H4	AI -> SRC	0.089	0.088	0.041	2.172	.030	Accepted
H5	MLB -> SRC	0.010	0.014	0.028	0.342	.733	Rejected
H6	AL_CW -> SRC	0.174	0.174	0.043	4.057	<.001	Accepted
H7	SRC -> OE	0.455	0.456	0.036	12.512	<.001	Accepted
H8	SRC -> SAT	0.229	0.229	0.043	5.381	<.001	Accepted
H9	AL_CW -> SAT	0.119	0.119	0.043	2.796	.005	Accepted
H10	IS -> SAT	0.261	0.261	0.042	6.237	<.001	Accepted
H11	AL_CW -> OE	0.230	0.230	0.039	5.860	<.001	Accepted
H12	AS -> SAT	0.232	0.233	0.041	5.667	<.001	Accepted

The bootstrapping results reveal that the proposed research model was robustly supported by the empirical data. Among the factors associated with scientific research competence (SRC), active learning engagement (PE) emerged as the strongest predictor (β = 0.279, $p < .001$), followed by assessment and feedback standardization (AS) (β = 0.229, $p <$ .001), course design quality and continuity (AL_CW) (β = 0.174, $p < .001$), and system and lecturer support (IS) (β = 0.140, $p = .001$). Meanwhile, AI use guidance linked with academic integrity (AI) was a statistically significant but relatively weak predictor of self-assessed SRC

(β = 0.089, p = .030, f^2 = 0.010). These findings suggest that the development of students' SRC relies not only on individual effort but also heavily on transparent assessment rubrics, logical curriculum sequencing, academic scaffolding, and institutional regulations regarding digital tools.

Crucially, Hypothesis H5, which predicted a negative relationship between methodological and academic language barriers (MLB) and SRC, was not supported. The relationship was found to be exceptionally weak and statistically non-significant (β = 0.010, t = 0.342, p = .733). In addition, SRC was positively and strongly associated with overall course effectiveness (OE) and student satisfaction (SAT). Moreover, institutional factors such as AL_CW, IS, and AS were also significantly related to the outcome variables, emphasizing the pivotal role of curriculum design, supervisory support, and assessment standardization in shaping the overall learning experience across the Project 1 and Project 2 sequence.

4.4. Summary of Quantitative Analysis Results

Table 9 summarizes the hypothesis testing results for the structural model. Out of the 12 proposed hypotheses, 11 were strongly supported by the empirical data ($p < .05$), confirming the theoretical validity and robustness of the model. The sole exception was H5, whose rejection provides vital nuances regarding students' coping behaviors, a phenomenon further explored through the qualitative findings.

4.5. Qualitative Data Analysis Results (In-depth Interviews)

To complement the quantitative findings, the study analyzed 19 semi-structured interviews. The thematic analysis was used to clarify how students experienced Project 1 and Project 2, why some relationships in the quantitative model appeared as they did, and what kinds of support students considered necessary for improving scientific research competence.

Overall, students described the Project sequence as academically valuable but mentally demanding. Most participants considered Project 2 more challenging than Project 1 because it required deeper methodological application, stronger theoretical reasoning, and more careful interpretation of research findings. Project 1 was often viewed as a starting framework that introduced students to questionnaire design, basic data collection, and the general structure of a research paper. However, several participants also reported that the connection between the two courses was not always smooth. Instead of developing into a clearly progressive learning sequence, Project 2 was sometimes perceived as repeating the structure of Project 1 with more detailed requirements. The transition from data collection to data analysis was especially difficult because students felt that they had not received enough preparation in SPSS or quantitative interpretation before being required to analyze and discuss results.

The interviews also showed that students developed certain aspects of SRC after completing the two courses, especially in searching for materials, designing questionnaires, organizing group work, and presenting research results. Nevertheless, their competence development was uneven. The most frequently mentioned bottlenecks were the Literature Review, Findings/Data Analysis, and Discussion sections. Students found it difficult to locate suitable academic sources, synthesize previous studies, identify a research gap, and transform statistical results into meaningful arguments. These difficulties were not only methodological but also linguistic. Many students reported that they had ideas but struggled to express them in an academic English style, paraphrase appropriately, and follow APA conventions. This confirms the "double burden" experienced by English Language students, who must learn both research methodology and academic language at the same time.

Several learning conditions helped explain why students' SRC varied across individuals and groups. Students who actively monitored deadlines, divided tasks clearly, searched for additional learning materials, and communicated regularly with teammates tended to understand the research process more deeply. Lecturer support also played an important role. Participants valued step-by-step guidance, direct correction, and specific formative feedback because these helped them recognize what was wrong and how to revise their work. In contrast, delayed or overly general comments made students feel uncertain and increased pressure. Teamwork was another influential factor. While a balanced and responsible group could reduce workload and support learning, unequal participation and free-rider behavior placed heavy pressure on group leaders and sometimes allowed less active members to pass the course without developing equivalent research competence.

The qualitative data also helped explain the unexpected quantitative result for methodological and academic language barriers (MLB). Although students described these barriers as serious, MLB did not show the hypothesized negative relationship with SRC in the structural model. One possible explanation is that the MLB scale may have captured students' awareness of research difficulties rather than only the actual obstructive level of those difficulties. Students with stronger research competence may be more capable of recognizing that Literature Review writing, SPSS analysis, and Discussion writing are difficult tasks. Therefore, they may report both higher SRC and higher awareness of barriers. In addition, students often used coping strategies to reduce the visible impact of barriers on their self-assessed competence. These strategies included relying on more capable group members, using templates or previous papers, and using AI tools for language support, idea generation, and data interpretation. As a result, the barriers remained important in students' lived experiences, but their direct negative association with SRC was not statistically confirmed.

AI use was common among the interviewed students. Participants reported using tools such as ChatGPT, Gemini, Claude, Grammarly, and QuillBot for brainstorming, grammar checking, paraphrasing, translation, summarizing documents, and sometimes interpreting data. Many students were able to distinguish between supportive and problematic AI use. They considered AI acceptable when it helped suggest ideas, check language, or clarify difficult concepts, but they regarded copying AI-generated content without understanding as a violation of academic integrity. However, because course-level AI guidelines were still unclear, students were unsure about the acceptable boundary of AI use. Under deadline pressure, some students admitted that AI could become a shortcut rather than a learning support tool, especially in writing the Findings and Discussion sections. This suggests that AI literacy and academic integrity guidance should be integrated more explicitly into the Project courses.

Finally, students' suggestions focused mainly on course design, learning materials, and feedback. Although the semester length of three to four months was considered generally sufficient for group work, the amount of knowledge was perceived as heavy and highly concentrated. Students expressed a strong need for sample papers, chapter checklists, templates, APA guidance, and clearer instructions on SPSS. Moodle was seen mainly as a submission platform rather than a learning resource repository. Participants suggested that Project 1 and Project 2 should be more clearly differentiated: Project 1 should focus on foundational research skills such as reading articles, identifying gaps, designing instruments, and practicing basic data analysis, while Project 2 should focus on applying those skills in a complete research paper. They also emphasized the need for timely, detailed, and constructive lecturer feedback throughout the course rather than only at the final submission stage.

5. Discussion

5.1. Discussion on the Scientific Research Competence of English Language Students

The mixed-methods findings show a multi-dimensional picture of HUFLIT students' SRC. Quantitatively, the self-assessment results were not concentrated only in the high group: 37.2% of students rated themselves as Good/Very good, while the remaining 62.8% were in the Average, Somewhat difficult, or Great difficulty groups. The qualitative data further explained the gap between perceived competence and real difficulties during the Project process.

In practice, students were often more confident in surface-level tasks such as basic literature searching, designing Google Form questionnaires, or making PowerPoint presentations. When they had to work on core tasks such as writing the Literature Review, analyzing data with SPSS, and writing the Discussion, many students admitted that they were confused and lacked direction. This gap may come from the fact that some students passed the course because of strong support from group leaders, unequal task division in groups, or heavy reliance on generative AI tools.

This reflects the "double burden" faced by English Language students (Ciraso-Cali et al., 2022; Le et al., 2023). They not only lack a strong background in statistical thinking, or the logic of quantitative data analysis, but are also limited by academic English when they need to paraphrase and argue according to international academic standards.

5.2. Discussion on Factors Associated with SRC

The structural model (PLS-SEM) supported 11 out of 12 research hypotheses. Because the study used cross-sectional self-report data, the following discussion interprets the paths as predictive relationships and associations rather than definitive causal effects. The combination of quantitative and qualitative data provides deeper explanations for these relationships.

5.2.1. The key role of learning engagement (PE) and lecturer support (IS).

The statistical results confirm that learning engagement (PE) was the strongest predictor of SRC (H1). The interview data reinforce this finding: students who actively monitored their progress, learned tools by themselves, and communicated well with their groups tended to develop a deeper understanding of the course. In addition, lecturer support (IS) worked as cognitive scaffolding. Lecturers who directly corrected students' work and provided detailed formative feedback helped students shape their research thinking, which is consistent with Self-Regulated Learning (Panadero & Broadbent, 2018).

5.2.2. The importance of assessment standardization (AS) and course design quality and continuity (AL_CW).

The study shows that rubrics and assessment transparency (AS) were positively related to students' competence and satisfaction. However, the qualitative data reveal a tension that deserves closer scrutiny from a curricular-design perspective. While the course sequence ostensibly follows a progressive model - Project 1 for data collection, Project 2 for methodology and full-report writing - the actual student experience suggests a structural misalignment that the current implementation has not resolved.

Specifically, students are required to produce a complete research report in Project 1, yet the statistical and methodological tools they need (particularly SPSS-based data analysis) are introduced substantively only in Project 2. This sequencing creates what Biggs (1996) would characterize as a failure of constructive alignment: the assessment task (a full report)

outpaces the instructional preparation. The consequence is predictable - students resort to compensatory strategies such as template reproduction, peer dependence, or AI-assisted writing to bridge the competence gap, rather than developing the intended research skills organically. In effect, the curriculum asks students to demonstrate competencies it has not yet taught.

This structural deficiency is not merely a scheduling inconvenience; it undermines the pedagogical rationale of the two-course sequence. If Project 1 is meant to build foundational skills, its assessment should be calibrated accordingly - focusing, for example, on instrument design, pilot data collection, and literature review drafting, rather than demanding a complete manuscript. Project 2 would then function as the integrative course where students apply the full research cycle, including data analysis and discussion writing, supported by explicit SPSS instruction introduced earlier or concurrently. The department's current arrangement, where the most analytically demanding assessment precedes the corresponding instruction, not only overloads students but also dilutes the diagnostic value of Project 1 grades as indicators of genuine research readiness.

5.2.3. The unexpected result of methodological and language barriers (MLB).

One of the more noteworthy findings of this study is the rejection of H5. The quantitative data show that methodological and academic language barriers (MLB) had a positive but statistically non-significant relationship with SRC ($\beta = 0.010$, $t = 0.342$, $p = .733$), even though the hypothesis predicted a negative relationship and the interview data clearly documented students' struggles with academic English, research methods, and data analysis.

This anomaly warrants careful methodological reflection rather than superficial dismissal. The most parsimonious explanation lies in a construct-measurement problem: the MLB scale, as operationalized, appears to tap into perceived difficulty awareness rather than actual obstructive impact. These are conceptually distinct constructs. A student with strong research competence may be more cognizant of the inherent challenges of SPSS analysis and academic writing precisely because they have engaged deeply enough with these tasks to appreciate their complexity. Conversely, a student who has relied heavily on group leaders or AI tools may report fewer barriers simply because they have not confronted them directly. The result is a paradoxical positive association - or, at best, a null one - between barrier perception and competence.

This interpretation is corroborated by the qualitative data. Students frequently described coping mechanisms - including delegation to more capable group members, reliance on AI-generated text, and reproduction of report templates - that effectively attenuated the impact of barriers on self-assessed SRC without eliminating the barriers themselves from the learning experience. In other words, the barriers remained real and substantial in students' narratives, but the compensatory behaviors severed the expected negative link to competence outcomes.

Future research should address this measurement limitation by distinguishing between two dimensions of barrier experience: (a) awareness (the extent to which students recognize methodological and linguistic challenges) and (b) obstruction (the degree to which those challenges actually impede research performance). A two-factor operationalization of MLB would likely produce a clearer - and theoretically consistent - picture of the barrier-competence relationship. This should be treated as a design limitation of the current study rather than evidence that barriers are unrelated to SRC.

5.3. Discussion on Artificial Intelligence (AI) and Academic Integrity

In the context of wide AI use in learning, this study views AI guidance and academic integrity as a construct that can positively contribute to SRC (H4) when managed through explicit institutional protocols. Interviews reveal that most students used AI in some capacity during their Project courses. When AI was deployed for legitimate purposes - brainstorming, grammar correction, structural suggestions - it helped students mitigate part of the linguistic burden that characterizes the EFL research context.

However, the boundary of integrity is becoming increasingly blurred in the absence of clear, course-level guidelines. Some students admitted that they relied too heavily on AI to write or explain content, especially in the Findings and Discussion sections, where original analytical reasoning is most critical. This creates risks on two fronts: first, the ethical risk of academic dishonesty; and second, the pedagogical risk that AI dependence replaces - rather than scaffolds - the development of critical thinking. The Project courses therefore need a formal and explicit set of AI use regulations, integrated into both the syllabus and the assessment rubric, so that students can deploy digital tools as support mechanisms rather than substitutes for the research thinking process.

6. Conclusion

6.1. General Conclusion of the Study

This study was conducted to evaluate the scientific research competence (SRC) of English Language students at HUFLIT through Project 1 and Project 2, and then propose a competence framework and improvement solutions. By applying a convergent mixed-methods design, combining PLS-SEM for 500 quantitative responses and thematic analysis for 19 semi-structured interviews, the study answered the initial research questions. Because the data were cross-sectional and self-reported, the findings should be understood as evidence of associations and predictive relationships rather than definitive causal effects.

The findings show that students' SRC developed through the two courses but not evenly. Positive progress was found in literature searching, questionnaire design, and basic academic thinking. However, major bottlenecks remained in quantitative data analysis using SPSS, argument writing in the Discussion section, and academic English.

Regarding the associated factors, the quantitative data show that 11 out of 12 hypotheses were supported. Active learning engagement (PE), course design quality and continuity (AL_CW), lecturer/system support (IS), assessment and feedback standardization (AS), and AI use guidance linked with academic integrity (AI) were positively associated with SRC. Among them, PE was the strongest predictor of SRC ($\beta = 0.279$, $p < .001$). In contrast, methodological and academic language barriers (MLB) showed a positive but statistically non-significant relationship with SRC ($\beta = 0.010$, $t = 0.342$, $p = .733$), although students still described them as important difficulties in interviews. This suggests that MLB may reflect students' awareness of difficulty and their coping strategies, not only the actual obstructive power of those barriers. In addition, SRC was positively associated with overall course effectiveness (OE) and satisfaction (SAT), confirming the central role of SRC in the course experience.

6.2. Proposed SRC Framework for English Language Students

Based on the theoretical foundation, course learning outcomes (CLOs), and practical findings at HUFLIT, the study proposes a Scientific Research Competence framework for English Language students. It includes five core competence groups:

1. Methodological Competence: The ability to identify a research gap related to linguistics, teaching methods, translation, or culture, and to choose a suitable research design, such as qualitative, quantitative, or mixed methods.

2. Academic Linguistic Competence: The ability to read and synthesize specialized English materials, use standard academic writing style, paraphrase appropriately, and follow citation rules strictly, such as APA 7th edition.

3. Data Analysis Competence: The ability to design unbiased questionnaires/interview questions, use software such as Excel and basic SPSS to process data, and interpret statistical results into logical arguments.

4. Collaborative and Managerial Competence: The ability to plan and manage the research timeline in a scientific way, communicate effectively, solve group conflicts, and reduce free-rider behavior.

5. Digital Integrity and AI Literacy: The ability to search reliable academic databases and use generative AI responsibly: using AI to support idea generation and grammar correction, but never misusing it to fake analytical arguments or plagiarize.

6.3. Recommendations for Improvement

Based on the findings and the proposed competence framework, the study offers the following solutions to improve the quality of the two Project courses:

For the Faculty and the University (course design improvement):

- Reposition the goals of Project 1 and Project 2: Reduce the overlap in requirements between the two courses. Project 1 should be structured mainly as a skills course, focusing on how to read articles, write a Literature Review, design questionnaires, and practice basic SPSS/Excel. Project 2 should be the application course in which students conduct a complete research paper.
- Issue "Guidelines for AI Use in Academic Work": Clear regulations should be provided in writing and in the rubric about the boundary between AI "support" and "plagiarism". Tasks should also require students to submit a prompt log to show their independent thinking.
- Build a shared learning material repository (LMS/Moodle): Moodle should be improved by providing standard templates, chapter checklists, and an APA handbook so that students can self-study more easily and reduce complete dependence on supervisors.

For supervisors:

- Strengthen formative assessment: Instead of only using final summative assessment, lecturers should create small submission milestones to provide timely feedback. Feedback should be constructive and show students how to revise, not only say whether something is right or wrong.
- Manage teamwork strictly: Groups should be required to prepare a clear Task Assignment Log from the beginning of the semester. An anonymous peer-assessment mechanism should be applied at the end of the course to individualize scores, ensure fairness, and reduce free-rider behavior.

For students:

- Students should improve their self-regulation and actively plan their research from the beginning of the semester instead of waiting until the deadline is very close.
- Students should strictly follow academic integrity principles and view scientific research as an opportunity to develop analytical thinking and critical thinking skills that directly support their future work and career development.

6.4. Limitations and Directions for Future Research

Although the study tried to ensure objectivity and scientific quality, it still has several limitations. First, the study was conducted at only one institution, Ho Chi Minh City University of Foreign Languages - Information Technology. Therefore, the generalizability of the findings to other universities with similar programs should be considered carefully. Second, the study used a cross-sectional design, reflecting students' competence at one point after they completed Project 1 and Project 2. It did not assess longitudinal changes in competence throughout the learning process. Third, the SRC variable was mainly measured through students' self-assessment. Therefore, future studies should include evidence from learning products, rubric scores, or independent lecturer assessment to verify actual competence. Fourth, as discussed in Section 5.2, the MLB construct may have conflated awareness of difficulty with actual obstruction by difficulty. Future research should operationalize these as separate sub-dimensions to more accurately capture the barrier-competence relationship and produce findings consistent with established theory (Ciraso-Cali et al., 2022; Le et al., 2023). Finally, because all quantitative variables were collected through the same survey, common method bias may still have influenced the observed relationships, even though the model showed acceptable reliability and validity.

In the future, further studies can expand the sample to different universities to increase representativeness. In addition, using an experimental design to directly test the effectiveness of applying the SRC framework, AI guidance handbook, or peer-review process in Project classes would provide more valuable and practical contributions to the teaching of scientific research methodology in Vietnam.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*(3), 347-364. https://doi.org/10.1007/BF00138871

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7-74. https://doi.org/10.1080/0969595980050102

Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (Eds.). (2018). *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*. Routledge. https://doi.org/10.4324/9781315109251

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, *3*, Article 22. https://doi.org/10.3389/feduc.2018.00022

Brookhart, S. M., & Chen, F. (2014). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343-368. https://doi.org/10.1080/00131911.2014.929565

Bui, T. L., & Tran, M. L. (2022). Các nhân tố ảnh hưởng đến ý định tham gia nghiên cứu khoa học của sinh viên Học viện Nông nghiệp Việt Nam [Factors associated with students' intention to participate in scientific research at the Vietnam National University of Agriculture]. *Tạp chí Khoa học Nông nghiệp Việt Nam*, *20*(11), 1550–1560. https://vie.vjas.vn/index.php/vjasvn/article/view/2154.

Ciraso-Cali, A., Martinez-Fernandez, J. R., Paris-Manas, G., Sanchez-Marti, A., & Garcia-Ravida, L. B. (2022). The research competence: Acquisition and development among undergraduates in education sciences. *Frontiers in Education*, *7*, Article 836165. https://doi.org/10.3389/feduc.2022.836165

Creswell, J. W., & Creswell, J. D. (2009). *Research design*. SAGE Publications Ltd.

Davidson, Z. E., & Palermo, C. (2015). Developing research competence in undergraduate students through hands on learning. *Journal of Biomedical Education*, *2015*, 1-9. https://doi.org/10.1155/2015/306380

Dawson, P. (2015). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347-360. https://doi.org/10.1080/02602938.2015.1111294

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: Students as partners in learning and teaching in higher education*. Higher Education Academy. https://www.advance-he.ac.uk/knowledge-hub/engagement-through-partnership-students-partners-learning-and-teaching-higher

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-014-0403-8

Le, T. T. T., Nguyen, T. T. H., Vu, N. Q. D., & Pham, H. H. (2025). Đề xuất khung năng lực khoa học mở cho các nhà nghiên cứu và thực hành giáo dục tại Việt Nam [Proposing an open science competence framework for educational researchers and practitioners in Vietnam]. *Tạp chí Giáo dục*, *25*(14), 1–6. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/3654

Panadero, E., & Broadbent, J. (2018). Developing evaluative judgement. In *Developing Evaluative Judgement in Higher Education* (pp. 81-89). Routledge. https://doi.org/10.4324/9781315109251-9

Prosekov, A. Y., Morozova, I. S., & Filatova, E. V. (2020). A case study of developing research competency in university students. *European Journal of Contemporary Education*, *9*(3), 592-602. https://doi.org/10.13187/ejced.2020.3.592

Reddy, Y. M., & Andrade, H. (2009). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435-448. https://doi.org/10.1080/02602930902862859

Sơn, H. Đ., & Mai, N. T. N. (2021). Các nhân tố ảnh hưởng đến sự tham gia nghiên cứu khoa học của sinh viên – nghiên cứu trường hợp đại học tài chính-marketing. *Tạp chí Nghiên cứu Tài chính - Marketing*, 13-24. https://doi.org/10.52932/jfm.vi49.92

Tính, T. T., Nhân, H. T., Khoa, L. V., Điền, T. T., Đức, P. M., Tìm, C. T., Nguyên, Đ. V., & Tấn, N. V. (2020). Hiện trạng và giải pháp nâng cao năng lực nghiên cứu khoa học của giảng viên trường đại học cần thơ. *Can Tho University Journal of Science*, *56*(4), 161. https://doi.org/10.22144/ctu.jvn.2020.095

Vo, T. M. N. (2023). Nghiên cứu về những yếu tố ảnh hưởng lên động cơ tham gia nghiên cứu khoa học của sinh viên [A study on the factors influencing students' motivation to participate in scientific research]. *Tạp chí Khoa học và Công nghệ Đại học Đà Nẵng*, *21*(4), 27–33. https://jst-ud.vn/jst-ud/article/view/8332

Xuan-Mai Le, H.-Y. P. (2023). *Impact of using analytic rubrics for peer assessment on EFL students' writing performance: an experimental study*. https://doi.org/10.5281/zenodo.7750831

Acknowledgment

The author sincerely thanks the Faculty of Foreign Languages, HUFLIT, the lecturers involved in teaching Project 1 and Project 2, and all student participants for their support, guidance, and valuable contributions to this study.

Bionote

Quoc Bao Mai is currently a student at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). His learning orientation focuses on research and practical application in the English Language major. He has a strong interest in modern educational methods and students' scientific research skills. This research is part of his effort to understand and improve training quality, aiming to build academic standards and develop students' research competence in the context of applying technology, especially AI, to learning.

#87 EFL STUDENTS' USE OF CHATGPT IN ARGUMENTATIVE ESSAY WRITING: INSIGHTS FROM A QUALITATIVE STUDY

Tran Nhu Uyen Le[1], Diem Bich Huyen Bui[2]

Abstract

The ability to construct well-supported arguments, address counterarguments, and provide effective refutations is essential in an argumentative essay. Despite increasing access to digital writing support tools, many English-majored students still struggle to produce well-developed argumentative essays. However, there is limited research on how students use ChatGPT to support argumentative writing and how they perceive its effectiveness. This study, therefore, investigates how 10 intermediate-level English majors use ChatGPT to support their argumentative writing through a qualitative design. The qualitative data revealed both positive and negative perspectives. On the one hand, ChatGPT was reported to play a supportive role in helping students strengthen their arguments by offering various argument options with relevant supporting details and evidence when prompted and reprompted correctly, and develop evaluative judgment to assess the effectiveness of the response and research skills to confirm the knowledge. It also supported students during the feedback stage by providing suggestions on language use and explanations of failures related to idea development, cohesion, and coherence in their writing. On the other hand, some students expressed concerns that without a clear understanding of ethical use and underlying cognitive processes, they may become overly reliant on the AI tool. Others reported limited prompting literacy and highlighted the need to develop this strategy to obtain more meaningful responses in a shorter time. These findings suggest that while AI tools can be helpful in improving argumentative writing, students still need guidance and should use them carefully to continue developing their own thinking and writing skills independently.

***Keywords*:** AI-assisted learning, argumentative writing, ChatGPT, writing support tool

1. Introduction

Argumentative writing is a central skill in English as a Foreign Language (EFL) academic contexts because it requires not only language accuracy but also higher-order abilities such as idea generation, logical reasoning, supporting claims with appropriate evidence, and engaging with counterarguments effectively (Song & Song, 2023). Research on academic writing has shown that EFL students often face difficulty with not only these aspects, but also other linguistic problems such as grammar, vocabulary, and organization (Sarıça & Gençoğlu, 2025). These can affect the quality of argumentative essays and lead to students' writing anxiety and low confidence (Song & Song, 2023). Oh and Hsieh (2025) also found that students' argumentative essays often contain vague thesis statements, underdeveloped arguments, insufficient evidence, weak transitions, unclear counterarguments, and grammatical inaccuracies. Given the challenges of argumentative writing and its crucial role in academic achievement, finding effective ways to support EFL learners in developing this essay genre is important in writing instruction.

[1] *BA, School of Languages, International University – Vietnam National University HCMC, Ho Chi Minh City, Vietnam*

[2] *Corresponding author, PhD Candidate, School of Languages, International University – Vietnam National University HCMC, Ho Chi Minh City, Vietnam; Email: bdbhuyen@hcmiu.edu.vn*

The rapid growth of artificial intelligence (AI) has provided opportunities for language learning. Among these, ChatGPT has attracted increasing attention because of its ability to generate text, provide instant feedback, rephrase ideas, and support students in improving their writing (Song & Song, 2023). Research suggests that when integrated into learning activities, ChatGPT can positively contribute to students' writing development, organization, and confidence in their writing abilities (Faiz et al., 2023). These features show that ChatGPT can support learners not only with language use but also with organizing and developing ideas in argumentative writing.

Despite these aforementioned benefits, there are still some challenges when using ChatGPT as a supporting tool for writing. Some studies suggest that AI tools may help students improve grammar and vocabulary, but these improvements are often limited to the surface level. Students may not engage deeply with critical thinking or argumentative reasoning, especially when instructional support is limited (Mai, 2026). Another concern is that students may rely too heavily on AI-generated suggestions without carefully evaluating the content. This can reduce learner autonomy and limit the development of original ideas (Sarıça & Gençoğlu, 2025). In addition, effective use of ChatGPT requires prompting literacy. Students need to write clear prompts, evaluate AI responses, and revise their queries when necessary. However, many learners have not yet fully developed these skills.

Although research on AI-assisted writing has grown rapidly, most of the existing literature has focused on writing outcomes, language accuracy, or general attitudes toward AI use. There is limited qualitative research on EFL students' perceptions of the usefulness and concerns of using ChatGPT for argumentative writing, especially in the Vietnamese higher education context. To address this gap, the present study investigates how intermediate-level English-major students perceive the benefits and challenges of using ChatGPT to support the writing of argumentative essays, based on participants' authentic practices and experiences. The findings are expected to contribute to EFL writing pedagogy by identifying both the advantages and potential difficulties of ChatGPT-supported argumentative writing. To fulfill this objective, the study seeks to answer the following research question*: How do EFL students perceive ChatGPT's role in supporting argumentative essay writing?*

2. Literature Review

2.1. Argumentative writing and its aspects

2.1.1. Argumentative writing

Argumentative writing is an important academic skill in higher education. It helps students not only demonstrate their language proficiency but also develop critical thinking, reasoning, and academic literacy (Ekalia et al., 2025). Ho et al. (2024) state that writing argumentative essays is demanding for EFL students because it requires language knowledge, idea generation, organization, grammar, vocabulary, mechanics, and writing style. Writers are also expected to consider and respond to opposing viewpoints in order to build a convincing argument (Walková & Bradfrod, 2022). For EFL students, this type of writing can be particularly challenging because they must manage both language use and argument development at the same time (Ekalia et al., 2025). As Ho et al. (2024) note, effective argumentative writing involves making a clear claim, providing supporting evidence, and addressing counterarguments to persuade readers.

2.1.2. Aspects of argumentative writing

One important aspect of argumentative writing is the development of a clear and arguable claim. The claim represents the writer's position on an issue and guides the overall direction of the essay. Effective argumentative writing requires students to address opposing viewpoints and provide rebuttals. According to Oh and Hsieh (2025), argumentative essays require students to present a clear thesis, develop supporting arguments, consider opposing views, and provide rebuttals. Similarly, Tai et al. (2025) and Darmawansah et al. (2024) highlight the importance of counterarguments and rebuttals in constructing strong arguments.

Another key aspect of argumentative writing is the use of evidence and reasoning. Effective arguments are supported by relevant facts, examples, and explanations rather than personal opinions alone. Evidence helps support the writer's claim, while reasoning explains how the evidence relates to the argument. Wang et al. (2024) note that strong arguments include elements such as claims, evidence, reasoning, and rebuttals. Besides, argumentative writing also requires critical thinking. Students need to analyze information, evaluate evidence, compare different viewpoints, and draw reasoned conclusions. As Suh et al. (2025) and Chen et al. (2025) suggest, argumentative writing and critical thinking are closely connected because students must examine information carefully and justify their positions with logical support.

Organization and coherence are also essential in argumentative writing. Ideas should be presented in a logical order and connected clearly throughout the essay. A typical argumentative essay includes an introduction that presents the issue and thesis statement, body paragraphs that develop arguments with evidence and reasoning, and a conclusion that emphasizes the writer's position. Tai et al. (2025) note that effective argumentation depends on both clear structure and appropriate language features, such as transitions and conclusion markers. These elements help readers follow and understand the argument more easily.

2. ChatGPT in supporting argumentative writing

Research suggests that EFL students use ChatGPT for brainstorming, organizing ideas, drafting, and revising their essays (Ho et al., 2024). Many students perceive ChatGPT as a helpful resource because it provides quick access to information, expands vocabulary, and assists with essay organization (Esmaeil et al., 2023; Annamalai et al., 2025). These features may help learners overcome difficulties in generating and developing arguments.

In argumentative writing, ChatGPT can also support the development of essay structure, counterarguments, and feedback. Studies have shown that students use ChatGPT to create outlines, organize arguments, and improve the coherence of their essays (Pratama et al., 2025). In addition, ChatGPT can provide immediate feedback on issues such as thesis statements, argument development, evidence use, cohesion, and language accuracy (Oh & Hsieh, 2025; Wang et al., 2024). Some research further suggests that ChatGPT may encourage students to consider alternative perspectives and develop stronger counterarguments, particularly when they interact with the tool through discussion-based or role-playing prompts (Ma et al., 2025). When used as a collaborative partner rather than a replacement for writing, ChatGPT may also promote reflection and critical thinking (Suh et al., 2025).

Despite these benefits, concerns have been raised about the use of ChatGPT in argumentative writing. Mai (2026) suggests that excessive reliance on AI-generated content may reduce students' critical engagement and lead to formulaic writing. In addition, ChatGPT may provide inaccurate, incomplete, or overly general responses, requiring students to verify information carefully (Chen et al., 2025). Researchers have also highlighted issues related to plagiarism, academic integrity, and the development of students' own rhetorical voice (Esmaeil

et al., 2023). These findings suggest that the effectiveness of ChatGPT depends largely on how critically and responsibly students use the tool.

Although previous studies have reported students' perceptions of ChatGPT in argumentative writing, existing research has primarily focused on the perceived benefits and challenges of the tool. For example, Esmaeil et al. (2023) found that students viewed ChatGPT as useful for idea generation, guidance, and writing support, while also expressing concerns about dependency, plagiarism, and reduced critical thinking. There is still relatively limited research about how EFL students actually use the tool while writing. To be more specific, it is unclear how students use ChatGPT to develop arguments, evaluate its responses, revise their essays, and check the accuracy of information. In addition, issues such as prompting skills, critical evaluation of AI-generated content, and the possibility of over-reliance have not been fully examined. Therefore, this study explores how intermediate-level EFL students use ChatGPT to support their argumentative essay writing and how they perceive its role in the writing process.

3. Methodology

3.1. Settings and participants

This study was conducted in an EFL context at a university, focusing on intermediate-level English majors. The participants were ten students, purposively selected based on their enrollment in a writing-intensive course and their experience with argumentative essay writing. All participants had similar English proficiency levels (CEFR B1–B2) to ensure consistency in their engagement with the ChatGPT tool.

3.2. Data collection

Data were collected using a qualitative method to capture the participants' experiences and perceptions (Ellis & Hart, 2023). Semi-structured interviews were conducted to explore participants' perceptions of ChatGPT's usefulness, the challenges they encountered, and their reflections on the tool's role in supporting their argumentative writing (Karatsareas, 2022).

3.3. Data analysis

All collected data, including interview transcripts, were analyzed using thematic analysis. The analysis focused on identifying patterns related to how students used ChatGPT to generate ideas, organize arguments, and incorporate counterarguments and evidence. The data were analyzed using thematic analysis. Interview recordings were first transcribed with the support of TurboScribe transcription software, and the transcripts were then checked with the original audio recordings to ensure accuracy.

The researchers generated initial codes, and grouped similar codes into major themes related to students' use of ChatGPT, such as idea generation, argument development, feedback use, prompting strategies, and concerns about over-reliance. To ensure the trustworthiness of the findings, both authors reviewed the coding and discussed any differences until they reached agreement (Cole, 2023). An expert in translation also reviewed the translated scripts. In addition, member checking was conducted to ensure the accuracy of the translations (Sahakyan, 2023).

4. Results

The interview data showed that students had both positive and cautious perceptions toward using ChatGPT in argumentative essay writing. Overall, they did not see ChatGPT as a replacement for their own thinking. Instead, they viewed it as a helpful tool that supported

different stages of writing, especially idea generation, developing arguments, revising, and improving language. Students also recognized some challenges, such as incorrect information, the risk of over-reliance, ethical concerns, and difficulties in writing effective prompts.

4.1. Strengthening arguments with ideas and evidence

Many participants reported that ChatGPT was particularly useful during the planning stage of argumentative writing. Students often described having difficulty generating ideas, especially when they were unfamiliar with the topic or could only think of one side of an issue. In these situations, ChatGPT helped them explore different perspectives, generate supporting details, and develop more balanced arguments.

Several students explained that ChatGPT provided multiple viewpoints rather than a single answer, allowing them to compare ideas before deciding which ones to include in their essays.

"It doesn't just give one answer, but multiple possible angles, so I can compare and choose the ones that fit my thinking" (P1).

Similarly, P3 explained that ChatGPT helped expand ideas when initial arguments were limited.

"Sometimes I only have one or two simple ideas, and I feel that they are not strong or diverse enough to develop into a full essay. So I use ChatGPT to generate more ideas or suggest different perspectives" (P3).

These comments suggest that students viewed ChatGPT as a brainstorming tool that broadened their thinking and helped them develop stronger arguments.

Another benefit frequently mentioned by participants was ChatGPT's ability to provide counterarguments. Students noted that argumentative writing requires considering different viewpoints, but they often focused only on their own position. ChatGPT helped address this issue by presenting alternative perspectives.

"It often provides both sides of an argument, including possible counterarguments, which helps me think more critically and avoid having a one-sided essay" (P3).

For some students, ChatGPT also reduced the pressure of getting started. Instead of facing a blank page, they could use AI-generated suggestions as a starting point for developing their own ideas.

"If I don't know how to start my essay or if I feel that my ideas are too limited, I ask ChatGPT to generate some initial ideas or give simple examples. This helps me reduce the pressure of starting" (P5).

Likewise, P9 highlighted the efficiency of receiving multiple ideas at once.

"It can provide multiple ideas at once, which saves time and helps me compare different options before choosing the most suitable ones" (P9).

Despite these benefits, participants emphasized that they did not simply copy ChatGPT's responses. Instead, they selected, modified, and combined AI-generated ideas with their own thinking.

"I usually don't take all the ideas from ChatGPT directly. Instead, I read through them and choose the ones that I think are relevant or suitable for my essay. Sometimes, I also combine its ideas with my own to create better arguments" (P6).

4.2. Developing evaluation through information verification

Although participants found ChatGPT useful, many emphasized that its responses should not be accepted without careful evaluation. Students were aware that ChatGPT could sometimes provide inaccurate, incomplete, or overly general information. As a result, they described a process of assessing AI-generated suggestions before deciding whether to use them in their essays.

Several participants explained that they did not simply copy ChatGPT's responses but actively evaluated their relevance and quality. For example, P1 stated:

> "I don't just copy its ideas. I usually read them, evaluate them, and then decide which ones are relevant to my argument" (P1).

A similar view was expressed by P10:

> "I also try to evaluate its suggestions instead of accepting them automatically" (P10).

These responses suggest that students maintained control over the writing process and treated ChatGPT as a source of suggestions rather than a source of final answers. Before incorporating AI-generated content into their essays, they considered whether the information was relevant, logical, and appropriate for their arguments.

Many reported that checking evidence, examples, and factual claims through other sources before using them in their writing is essential.

As P1 explained:

"I don't fully trust ChatGPT, especially when it provides evidence or examples. Sometimes the information is not accurate, so I always check external sources like Google or news websites" (P1).

Similarly, P3 noted:

> "For important ideas, critical thinking, or evidence, I still prefer to use my own thinking or search for information from reliable sources like Google or academic websites" (P3).

These comments indicate that participants were aware of the limitations of AI-generated information and recognized the importance of using reliable sources. Rather than treating ChatGPT as an authority, they used it as a starting point and then verified important information independently.

Interestingly, some students felt that this process encouraged them to become more critical when working with information. Because they knew that ChatGPT could make mistakes, they became more cautious when selecting evidence and supporting details for their arguments. In this sense, ChatGPT not only supported writing but also encouraged students to evaluate, question, and verify information more carefully.

Overall, participants reported that using ChatGPT involved more than simply receiving information. It required them to evaluate the quality of AI-generated responses, decide which suggestions were useful, and verify information through external sources before incorporating it into their writing.

4.3. Supporting revision through feedback and language improvement

Another benefit mentioned by participants was the use of ChatGPT during the revision stage. Many students reported using the tool after completing a first draft to receive feedback on grammar, vocabulary, sentence structure, academic tone, coherence, and overall essay

organization. For several participants, this was one of the most valuable ways ChatGPT supported their writing.

Some students explained that the quality of feedback improved when they provided specific criteria. For example, P2 shared:

> "After I finish writing my essay, I usually input my writing together with the rubric or scoring criteria, and I ask ChatGPT to evaluate my essay based on those criteria" (P2).

Similarly, P7 described assigning ChatGPT a specific role when requesting feedback:

> "I assign ChatGPT a role, such as an IELTS examiner or a lecturer, and include specific criteria like coherence, vocabulary, grammar, and task response. When I do this, the feedback becomes much more detailed, structured, and relevant to my needs" (P7).

These comments suggest that students used ChatGPT not only to identify problems in their writing but also to obtain more targeted feedback that aligned with assessment expectations.

Participants particularly valued feedback related to language use. Several students reported that ChatGPT helped them identify grammar and sentence structure errors that they often overlooked when revising independently.

As P4 explained:

> "After I finish writing my essay, I use ChatGPT to check grammar, spelling, and sentence structure. It helps me identify many small mistakes that I cannot notice by myself" (P4).

In addition to correcting errors, ChatGPT also helped students improve the academic quality of their writing. P6 noted:

> "Sometimes my writing is too informal, and ChatGPT suggests more formal expressions that are more suitable for academic writing or IELTS essays" (P6).

Likewise, P4 highlighted the role of ChatGPT in vocabulary improvement:

> "One of my biggest problems in writing is repetition. I often use the same words again and again without realizing it. ChatGPT helps me replace those words with synonyms and suggests more varied expressions" (P4).

These responses indicate that students viewed ChatGPT as a useful tool for improving language accuracy, vocabulary variety, and academic style.

Beyond language correction, several participants felt that ChatGPT helped them better understand their weaknesses as writers. Rather than simply providing corrections, the tool sometimes encouraged reflection on recurring problems.

For example, P1 stated:

> "One important experience I had is that ChatGPT helps me become more aware of my weaknesses. For example, I realized that I often write unclear topic sentences or repeat the same vocabulary" (P1).

Similarly, P5 appreciated the explanations provided by the tool:

> "One thing I find particularly useful is that ChatGPT not only corrects the mistakes but also explains why they are wrong. This helps me understand my errors more deeply and avoid repeating them in future writing" (P5).

These comments suggest that students valued feedback that supported learning rather than simple error correction.

Participants also reported using ChatGPT to improve idea development, cohesion, and essay organization. According to P10:

> "ChatGPT helps me understand how to structure an argumentative essay more clearly. For example, it shows me how to write a clearer introduction, how to develop body paragraphs with better topic sentences and supporting ideas, and how to write a more effective conclusion" (P10).

Likewise, P5 explained:

"Before using ChatGPT, I sometimes confused counterarguments and rebuttals, and my ideas were not clearly structured. ChatGPT helps me by suggesting supporting ideas, giving examples, and showing different ways to present arguments more effectively" (P5).

These responses indicate that ChatGPT supported not only linguistic accuracy but also students' understanding of how argumentative essays should be organized and developed.

4.4. Concerns about over-reliance, ethical use, and prompting skills

Although participants generally viewed ChatGPT positively, they also expressed concerns about its potential drawbacks. The most frequently discussed issue was over-reliance. Several students worried that depending too heavily on ChatGPT, particularly during idea generation, could reduce their ability to think independently and develop arguments on their own. Another issue frequently mentioned by participants was the need to verify information. Students recognized that although ChatGPT could generate convincing responses, the information was not always accurate.

Some participants reflected on how their use of ChatGPT changed over time. For example, P8 explained that after frequently relying on ChatGPT for brainstorming, they became aware of their growing dependence on the tool and intentionally adjusted their writing process.

> "At the beginning, I used ChatGPT quite frequently, especially in the brainstorming stage, because it was very easy to generate ideas quickly. However, after using it for some time, I realized that I was becoming too dependent on it" (P8).

As a result, P8 reported writing independently first and using ChatGPT mainly during the revision stage.

> "Now, I try to write my essay by myself first, without using ChatGPT in the early stages. Only after that do I use ChatGPT, mainly in the revision stage" (P8).

Similar concerns were raised by other participants. Some believed that excessive use of ChatGPT could weaken students' reasoning and idea-generation abilities.

> "If we depend on it for everything, especially for generating ideas, we may lose the ability to think independently" (P6).

> "If we rely on it too much, especially for generating ideas or writing entire essays, it can make us lazy and reduce our ability to think independently" (P10).

These comments suggest that students were aware of the need to maintain their own role in the writing process rather than allowing ChatGPT to do the thinking for them.

Another issue discussed by participants was the ethical use of ChatGPT. Students generally agreed that the tool should support learning rather than replace their own work. Many emphasized the importance of maintaining authorship and ensuring that the final essay reflected their own ideas and understanding.

For example, P5 explained that they did not use ChatGPT responses directly but instead adapted them to fit their own arguments.

> "I usually read the ideas it suggests, then select or adapt the ones that fit my understanding instead of copying them directly. In this way, I feel that I can still maintain my own voice in writing" (P5).

Likewise, P8 emphasized the importance of seeing ChatGPT as a learning resource.

> "I try to limit my usage and use ChatGPT more as a learning tool rather than a shortcut" (P8).

These responses indicate that participants were conscious of the ethical challenges associated with AI-assisted writing and attempted to use the tool responsibly.

A third concern involved prompting skills. Several students explained that the usefulness of ChatGPT depended heavily on the quality of the prompts they provided. Participants noted that vague prompts often produced generic responses, whereas detailed prompts resulted in more relevant and useful feedback.

As P2 explained:

> "When I write a clear and detailed prompt, the feedback is very useful. But when my prompt is too general, the feedback becomes vague and not helpful. So I think prompt-writing is a very important skill" (P2).

Similarly, P7 highlighted the importance of providing context and specific instructions when seeking feedback.

> "When my prompt is clear, specific, and includes context, the feedback is very useful and even quite similar to teacher feedback" (P7).

Some participants also described that they often needed to revise their prompts and ask follow-up questions to obtain more meaningful responses.

> "Sometimes the feedback is too general, especially when my prompt is not clear or detailed enough. In those cases, I need to ask follow-up questions or refine my prompt to get better answers" (P5).

5. Discussion

5.1. Strengthening arguments with ideas and evidence

A key finding of this study is that students viewed ChatGPT as a useful tool during the planning stage of argumentative writing. Many participants reported using ChatGPT when they had difficulty generating ideas, lacked background knowledge, or did not know how to begin an essay. As shown in the interview data, students emphasized ChatGPT's ability to provide multiple perspectives, supporting details, examples, and counterarguments. Rather than accepting a single response, they often compared different suggestions and selected those that best matched their own viewpoints. This finding suggests that students used ChatGPT not simply to obtain content but to expand and develop their ideas before writing.

This is consistent with Esmaeil et al. (2023) and Ho et al. (2024), who argued that ChatGPT can support idea generation, outline preparation, and content development in argumentative writing. Similarly, Darmawansah et al. (2024) found that ChatGPT-supported argumentation helped students strengthen their arguments through reasoning, feedback, and counterarguments. The present study supports these findings and further shows how students personally experienced this support. For many participants, ChatGPT reduced the difficulty of getting started and helped them explore different sides of an issue. This is important in argumentative writing because students are expected to consider diverse perspectives before developing their own position.

5.2. Developing evaluation through information verification

Another finding of this study is that students did not accept ChatGPT's responses immediately. Although participants generally found the tool helpful, they emphasized the importance of evaluating its suggestions before using them in their essays. Several students reported reading, comparing, selecting, adapting, or even rejecting ChatGPT-generated responses. Many also stated that they checked evidence and information using Google, news websites, or academic sources. These responses show that students were aware of the limitations of AI-generated content and tried to use the tool critically.

This result supports Chen et al. (2025)'s argument that the value of ChatGPT depends on how students interact with it. Rather than using ChatGPT passively, participants in this study engaged in a process of evaluation and verification. Interestingly, awareness of ChatGPT's limitations appeared to make some students more cautious when selecting information. As a result, using ChatGPT may encourage students to think more carefully about the relevance and reliability of information before including it in their writing. This indiates that ChatGPT can support the development of evaluative judgment when students are aware that its responses are not always accurate.

5.3. Supporting revision through feedback and language improvement

Qualitative data also showed that students used ChatGPT most frequently during the revision stage. Participants often described using the tool after completing a first draft to receive feedback on grammar, vocabulary, sentence structure, academic tone, coherence, and essay organization. Several students also reported providing rubrics or assigning specific roles, such as IELTS examiner or lecturer, to obtain more focused feedback. These practices indicate that students actively guided the feedback process to meet their own writing needs.

This finding is in line with Oh and Hsieh (2025), who found that ChatGPT can provide immediate feedback on both language and argumentative features. Similarly, Pratama et al. (2025) reported that students viewed AI-generated feedback positively because it was accessible and responsive. Consistent with these studies, participants in the present study valued ChatGPT because it allowed them to revise repeatedly and receive feedback whenever they needed it. More importantly, students reported that ChatGPT helped them identify recurring weaknesses, understand why certain errors occurred, and improve both language use and argument development. This suggests that ChatGPT can support not only revision but also students' awareness of their own writing strengths and weaknesses.

5.4. Concerns about over-reliance, ethical use, and prompting skills

However, students also reported several concerns about using ChatGPT. The most common concern was over-reliance. Several participants worried that frequent use of ChatGPT, particularly for idea generation, could reduce their ability to think independently. Some

students even described changing their writing habits after realizing that they had become too dependent on the tool. Instead of using ChatGPT from the beginning, they started generating ideas on their own and used the tool mainly for revision.

This finding is consistent with Mai (2026), who found that excessive reliance on ChatGPT may result in formulaic writing and lower levels of critical engagement. Similarly, Esmaeil et al. (2023) reported concerns that dependence on ChatGPT could weaken students' critical thinking skills. The findings of the present study support these concerns. At the same time, they show that some students were aware of this risk and attempted to use ChatGPT in a more balanced way. Rather than replacing their own thinking, they viewed ChatGPT as a tool that should support the writing process.

Students also emphasized the importance of writing effective prompts. Many participants explained that the quality of ChatGPT's responses depended heavily on the quality of the prompts they provided. Detailed prompts usually resulted in more useful and relevant feedback, while vague prompts often led to general or less helpful responses.

This result supports Ma et al. (2025), who emphasized the role of carefully designed prompts in obtaining useful support for argumentative writing. The findings of the present study suggest that prompting should be viewed as an important part of AI literacy. Students need to know how to provide context, specify their needs, ask follow-up questions, and refine their prompts when necessary. Therefore, if ChatGPT is to be integrated into EFL writing instruction, teachers should provide guidance not only on how to use AI tools but also on how to interact with them effectively.

6. Conclusion

This study examined how intermediate-level EFL students used ChatGPT to support their argumentative essay writing. The findings suggest that ChatGPT can support different stages of the writing process, particularly idea generation, argument development, and revision. Participants also reported using ChatGPT to explore alternative perspectives, develop counterarguments, and improve language use. Importantly, many students described evaluating and adapting AI-generated responses rather than accepting them directly, suggesting that ChatGPT can support critical engagement when used appropriately. However, the study also highlights several concerns, including over-reliance, ethical issues, and limited prompting literacy. Students may benefit less from ChatGPT if they use it passively or depend on it to replace their own thinking. Therefore, ChatGPT should be integrated into EFL writing instruction with clear guidance on responsible use, source verification, and effective prompting.

Although this study contributes to the understanding of students' experiences with ChatGPT in argumentative writing, there are still some limitations. The study involved a relatively small group of students from a single educational context and relied primarily on interview data. Future research could build on these findings by examining a larger and more diverse sample of learners. In particular, experimental or longitudinal studies could investigate whether and to what extent ChatGPT use leads to measurable improvements in argumentative writing quality over time.

Acknowledgments

The authors would like to thank all participants for their time and valuable contributions to this study. Their participation was essential to the completion of this research. We also sincerely

appreciate the anonymous reviewers for their insightful comments and constructive feedback, which greatly helped improve the quality and clarity of this paper.

References

Annamalai, N., Uthayakumaran, A., Bervell, B., & Kumar, R. (2025). Examining the use of ChatGPT in argumentative writing: A mixed methods study. *Journal of Advanced Academics, 36*(3), 521-550. https://doi.org/10.1177/1932202X251333859

Chen, X., Jia, B., Peng, X., Zhao, H., Yao, J., Wang, Z., & Zhu, S. (2025). Effects of ChatGPT and argument map (AM)-supported online argumentation on college students' critical thinking skills and perceptions. *Education and Information Technologies, 30*(12), 17623-17658. https://doi.org/10.1007/s10639-025-13471-2

Cole, R. (2023). Inter-rater reliability methods in qualitative case study research. *Sociological Methods & Research, 53*(4), 1944-1975. https://doi.org/10.1177/00491241231156971

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G.-J. (2024). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies, 30*(3), 3803-3827. https://doi.org/10.1007/s10639-024-12986-4

Ekalia, Y. J., Jemadi, F., & Susanto, I. (2025). Critical thinking skills and argumentative writing ability: Is there any correlation? *DIAJAR: Jurnal Pendidikan Dan Pembelajaran, 4*(3), 471-482. https://doi.org/10.54259/diajar.v4i3.5108

Ellis, J. L., & Hart, D. L. (2023). Strengthening the choice for a generic qualitative research design. *The Qualitative Report, 28*(6), 1759-1768. https://doi.org/10.46743/2160-3715/2023.5474

Esmaeil, A. A. A., Kiflee@Dzulkifli, D. N. A., Maakip, I., Mantaluk, O. O., & Marshall, S. (2023). Understanding student perception regarding the use of chatgpt in their argumentative writing: a qualitative inquiry. *Jurnal Komunikasi: Malaysian Journal of Communication*, 39(4), 150-165. https://doi.org/10.17576/jkmjc-2023-3904-08

Faiz, R., Bilal, H. A., Nawaz, S., & Asghar, I. (2023). Optimizing ChatGPT as a writing aid for EFL learners: Balancing assistance and skill development in writing proficiency. *Linguistic Forum - A Journal of Linguistics, 5*(3), 24-37.

Karatsareas, P. (2022). Semi-structured interviews. In R. Kircher & L. Zipp (Eds.), *Research methods in language attitudes* (pp. 99–113). Cambridge University Press. https://doi.org/10.1017/9781108867788.010

Ma, Z.-Q., Cui, X., Liu, W.-P., Tu, Y.-F., & Hwang, G.-J. (2025). ChatGPT-assisted collaborative argumentation: Effects of role-playing prompts on students' argumentation outcomes, processes, and perceptions. *Educational Technology & Society, 28*(3), 133-150. https://doi.org/10.30191/ETS.202507_28(3).SP09

Mai, H. T. (2026). Evaluating the impact of using chatgpt on EFL students' argumentative skills. Discover Education, 5(1). https://doi.org/10.1007/s44217-026-01153-1

Ho, H. B. N., Ho, H. B. N., & Thai, C. D. (2024). EFL students' perceptions and practices of using ChatGPT for developing English argumentative essay writing skills. *European Journal of Alternative Education Studies, 9*(1) 168-216. https://doi.org/10.46827/ejae.v9i1.5341

Oh, H.-J., & Hsieh, S.-F. (2025). An analysis of chatgpt-generated feedback on argumentative essays in korean EFL classrooms. Modern English Education, 26, 166-181. https://doi.org/10.18095/meeso.2025.26.1.16

Pratama, R. D., Widiati, U., & Hakim, L. N. (2025). Effects of human-ai collaborative writing strategy on EFL students' argumentative writing skills. Computer Assisted Language Learning, 1-32. https://doi.org/10.1080/09588221.2025.2503900

Sahakyan, T. (2023). Member-checking through diagrammatic elicitation: Constructing meaning with participants. *TESOL Quarterly, 57*(2), 686-701.

Sarıça, T. & Gençoğlu, S. D. (2025). EFL students' perceptions of AI-assisted writing tools: A systematic narrative hybrid review. *The Literacy Trek, 11*(1), 1–31. https://doi.org/10.47216/literacytrek.1669804

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, 1-14. https://doi.org/10.3389/fpsyg.2023.1260843

Suh, S., Bang, J., & Han, J. W. (2025). Developing critical thinking in second language learners: Exploring generative AI like ChatGPT as a tool for argumentative essay writing. *arXiv*, 1-23. https://arxiv.org/abs/2503.17013

Tai, H. Y., Lin, M. F., & Chen, Y. S. (2025). Incorporating ChatGPT into genre-based instruction for argumentative writing among EFL college students. *International Journal of Applied Linguistics, 36*(1), 767-781. https://doi.org/10.1111/ijal.12777

Walková, M., & Bradford, J. (2022). *Constructing an argument in academic writing across disciplines*. *ESP Today, 10*(1), 22–42. https://doi.org/10.18485/esptoday.2022.10.1.2

Wang, L., Chen, X., Wang, C., Xu, L., Shadiev, R., & Li, Y. (2024). ChatGPT's capabilities in providing feedback on undergraduate students' argumentation: A case study. *Thinking Skills and Creativity, 51*, 1-14. https://doi.org/10.1016/j.tsc.2023.101440

Bionote

Tran Nhu Uyen Le is a graduate majoring in English Linguistics and Literature at International University, VNU-HCM, Ho Chi Minh City, Vietnam. She works as an English teacher for both young learners and adults and has a strong commitment to academic research. Her research interests lie in phonetics and phonology, technology-enhanced language learning (TELL), and 21st-century learning skills.

Email: nhuuyen170404@gmail.com

Diem Bich Huyen Bui is currently a full-time lecturer at the School of Languages, International University, VNU-HCM, Ho Chi Minh City, Vietnam. She completed her undergraduate studies in TESOL and earned her MA in Applied Linguistics from La Trobe University, Australia. She is currently pursuing her PhD in TESOL. Her research interests include formative assessment, learner autonomy, skills development, and technology-enhanced language learning (TELL).

Email: bdbhuyen@hcmiu.edu.vn

#89 EFL TEACHERS' PERCEPTIONS OF AI-SUPPORTED AND DIGITAL TECHNOLOGIES IN IELTS TEACHING FOR HIGH-QUALITY STUDENTS

Thai, Ho-Van[1]

Abstract

AI and Digital technologies have been taking a crucial focus in the English language learning in higher education in high-stakes preparation institutions like IELTS. While there are many international studies that consider the role of technology in facilitating language learning, fewer have included examining how Vietnamese EFL teachers perceive the factors influencing their adoption of digital and AI-driven tools in teaching high-quality students to take the IELTS. This research investigates what influences the adoption of such technologies by teachers in Vietnamese higher education institutions which provide high-quality programmes. A mixed-method approach - survey responses from 50 IELTS lecturers together with interviews were utilised to elicit more detailed findings. Results indicate that teachers' willingness to use AI-supported and digital technologies is driven by five main factors: pedagogic beliefs, technological knowledge and digital competence, personal interest and perceived usefulness, time constraints and workload, as well as institutional culture, infrastructure, and policy support. Despite the gains that teachers cited that AI enhances performance such as automatic feedback, enhanced learner engagement, accelerated test-prep, teachers recognised concern about AI effectiveness and reliability and students' dependence upon AI tools and possible inconsistencies with IELTS assessment criteria. These findings highlight the necessity of systematic AI literacy training, clearer institutional guidelines & implementation of digital pedagogy principles in the curriculum and policy framework for the more critical and potent use of AI-supported technologies in IELTS-directed instruction for students.

***Keywords*:** EFL, Artificial Intelligence, perceptions of AI, IELTS, high-quality

1. Introduction

1.1. Background

The rapid evolution of generative artificial intelligence (AI) and specialized digital applications has radically reshaped the landscape of higher education globally, and English as a Foreign Language (EFL) instruction is no exception (Davis, 2024; Kamalov et al., 2023). Within the context of high-stakes language assessment, particularly the International English Language Testing System (IELTS), digital and AI-driven tools have transitioned from peripheral supplementary materials to core pedagogical infrastructure. In Vietnamese higher education, this technological shift intersects with institutional efforts to elevate academic standards through the implementation of "high-quality" academic programs. These specialized programs explicitly demand superior English proficiency from undergraduates, frequently utilizing ambitious IELTS band scores (ranging from 6.5 to 7.5 or higher) as strict benchmarks for global academic integration and graduation clearance. Consequently, tertiary EFL lecturers

[1]*College of Foreign Economic Relations, Ho Chi Minh City, Vietnam; Email: hvthai@cofer.edu.vn*

must navigate the intricate demands of high-stakes exam preparation while simultaneously integrating complex digital ecosystems into their daily instructional routines.

1.2. Statement of the problem

While technology acceptance in EFL is a widely researched domain, the integration of AI tools in specialized, high-stakes testing contexts presents a unique, underexplored pedagogical crisis. IELTS preparation is inherently challenging due to its rigorous assessment rubrics, which require candidates to demonstrate advanced cognitive skills and highly structured academic writing. The core problem emerges when these generalized AI tools intersect with highly pragmatic learner cohorts, specifically, logistics-majored students enrolled in high-quality English programs.

Unlike traditional language or business English majors, logistics students are rigorously trained to prioritize operational efficiency, problem-solving, and task completion. When preparing for the IELTS, this pragmatic mindset often drives them to utilize generative AI not as a cognitive scaffold, but as an academic shortcut to outsource language production, creating a severe algorithmic dependency cycle. Despite the growing ubiquity of automated writing evaluation and generative AI, there is a critical empirical gap regarding how Vietnamese EFL lecturers perceive and manage this specific crisis. Educators are currently operating in an institutional policy vacuum, struggling to balance the efficiency gains of AI with the urgent need to mitigate student overreliance and maintain alignment with strict IELTS band descriptors. The inquiry is guided by the following research questions:

RQ1. *What are the primary factors influencing EFL teachers' perceptions of AI-supported technologies in IELTS teaching?*

RQ2. *How do institutional constraints and the specific pragmatic learning behaviors of non-language majors impact teachers' instructional agency regarding AI integration?*

Drawing upon the Technology Acceptance Model (TAM) and the Technological Pedagogical Content Knowledge (TPACK) framework, this study proposes the following quantitative hypotheses:

H1. *Perceived usefulness (PU) and TPACK positively influence teachers' willingness to integrate AI-supported tools in IELTS preparation.*

H2. *Time constraints, heavy workload, and lack of institutional policy support negatively moderate the pedagogical integration of AI-driven tools.*

H3. *Teachers' concerns regarding AI reliability and student algorithmic dependence negatively impact their unconditional acceptance of generative AI in high-stakes writing instruction.*

1.4. Delimitations

To ensure empirical depth and contextual relevance, this study is delimited in its scope. Geographically and institutionally, the research targets higher education institutions located in Ho Chi Minh City, Vietnam. Demographically, the participant pool is limited to 50 tertiary EFL lecturers explicitly tasked with teaching IELTS preparation courses. Furthermore, the instructional context is confined to high-quality undergraduate programs, specifically targeting cohorts of logistics-majored students learning English for specific professional and academic purposes. This study explicitly excludes English linguistics majors.

2. Literature Review

2.1. Theoretical perspectives on AI-supported instruction in IELTS teaching

The growing presence of AI in educational settings has prompted renewed attention to how emerging technologies shape teaching practices (Dimitriadou & Lanitis, 2023). In English language education, AI-supported instruction refers to the use of intelligent systems, such as automated feedback tools, intelligent tutoring applications, and generative language models, to support learners during the language acquisition process (Liu et al., 2025).

Understanding teachers' engagement with such tools in high-stakes testing contexts like IELTS requires a theoretically grounded perspective that recognizes teaching as a cognitively and socially situated activity (Poquet, 2024). Teacher cognition theory provides a useful foundation for examining how teachers interpret and respond to instructional innovations, asserting that teachers' beliefs, perceptions, and prior experiences strongly influence classroom decision-making and instructional behavior (Bakshi, 2021). The introduction of AI into IELTS instruction, therefore, does not lead to uniform pedagogical change; rather, it is mediated by teachers' perceptions of AI regarding its instructional value, risks, and legitimacy (Wu et al., 2026).

In addition, the TAM suggests that individuals' adoption of new technologies is influenced by perceived usefulness and perceived ease of use (Liesa-Orús et al., 2023). Within high-stakes preparation contexts, these perceptions are closely linked to pedagogical beliefs and professional identity (Suarez & McGrath, 2022). Complementing these perspectives, Technological Pedagogical Content Knowledge (TPACK) theory emphasizes that meaningful technology integration depends on teachers' ability to align technological knowledge with pedagogical strategies and subject matter expertise, specifically the rigid evaluation criteria of the IELTS band descriptors (Zhang & Li, 2025).

2.2. Teachers' perceptions of AI-supported in IELTS teaching

Teachers' perceptions of AI-supported tools represent a critical cognitive driver that shapes instructional readiness (Wu et al., 2026). In high-stakes IELTS training, teachers exhibit a complex mixture of optimism and caution. Educators often perceive AI tools as highly useful for enhancing feedback quality, supporting continuous drafting, and managing the heavy instructional workload associated with grading writing and speaking components (An et al., 2022; Bower et al., 2024). When AI is perceived as an administrative asset that improves performance, teachers are more likely to integrate it as a resource that extends their professional role (Wu et al., 2026).

On the other hand, serious concerns persist regarding the reliability and alignment of AI feedback with official IELTS evaluation criteria (Li & Huang, 2025). Anxiety regarding academic integrity, potential algorithmic bias, and the fear of reducing students' critical thinking through overreliance often diminishes teachers' acceptance (Gustilo et al., 2024; Li & Huang, 2025). Therefore, teachers' perception does not function as a simple binary attitude but as an evaluative interpretive framework through which they negotiate the boundaries between automated efficiency and pedagogical validity (Azis & Numan, 2026).

2.3. Professional development and teacher AI competence

Professional development (PD) serves as the primary capability-building mechanism through which teachers construct technological knowledge, negotiate professional norms, and develop shared digital pedagogy principles within communities of practice (Chiu et al., 2025;

Dlamini et al., 2024). In the rapidly evolving domain of AI, traditional technical training that focuses solely on software proficiency is insufficient (Azis & Numan, 2026).

Effective PD must address both pedagogical implementation and ethical awareness, equipping IELTS instructors with the critical literacy needed to distinguish accurate automated outputs from algorithmic errors (Mouta et al., 2025). Without continuous institutional learning support, teachers' positive perceptions may fail to manifest as actual classroom innovation. PD effectively bridges the gap between perception and practice by enhancing instructional competence and strengthening teachers' confidence to exercise professional agency when integrating AI tools into high-stakes test environments (Wu et al., 2026).

2.4. Institutional and ethical support in higher education

Institutional and ethical support (IES) reflects the macro-environmental conditions, including organizational norms, structural infrastructure, and explicit policy frameworks, that regulate and legitimize educational innovations (Balzano et al., 2024). According to institutional theory, the enactment of individual capacity is highly contingent upon the stability of organizational structures (Wu et al., 2026).

When universities provide clear, explicit guidelines regarding the ethical boundaries of generative AI and intellectual property, teachers experience lower perceived risks and higher professional safety (Holmes & Porayska-Pomsta, 2023). Conversely, ambiguous or weak institutional governance forces instructors to rely entirely on personal risk management, which can severely constrain the deployment of innovative practices (Mouta et al., 2025). In early technological adoption phases, IES can function as a crucial compensatory or stabilizing structure that reduces variability driven by differences in individual teacher preparedness (Wu et al., 2026).

2.5. Pedagogical integration of AI and washback on IELTS rubrics

Pedagogical integration (PIAI) denotes the extent to which digital and AI-driven tools are purposefully embedded into instructional designs, daily classroom interactions, and assessment paradigms (Wu et al., 2026). In an exam-driven curriculum like IELTS, integration is intensely dominated by the washback effect, the direct impact of testing formats on teaching methodologies (Latif & Wasim, 2022; Lei & Lei, 2025). Meaningful PIAI requires that automated evaluation tools accurately reflect the hyper-specific public band descriptors of the exam, including lexical resource, grammatical range, and coherence (Wu et al., 2026). Effective pedagogical integration occurs when AI functions as a scaffold for independent learner development, rather than a shortcut that automates language production. Achieving this balance requires constant professional calibration to ensure that computer-assisted assessment practices align with authentic human-scored language standards (Azis & Numan, 2026).

3. Methodology

3.1. Research design

This study used a convergent parallel mixed-methods design to investigate tertiary EFL lecturers' perceptions and experiences regarding AI integration. This design allowed quantitative survey data and qualitative interview narratives to be collected concurrently and analyzed iteratively, facilitating the triangulation of findings and enriching the interpretation of the factors affecting technology adoption (Abedi et al., 2024; Denny & Weckesser, 2022). The quantitative component mapped broad institutional and pedagogical trends using descriptive statistics, while the qualitative component captured the nuanced, context-dependent

reasoning underlying teachers' decisions to integrate, modify, or limit specific tools (Hall & Liebenberg, 2024).

3.2. Participants and Context

The participant sample comprised 50 tertiary EFL lecturers (N=50) actively teaching IELTS preparation modules within high-quality undergraduate tracks across higher education institutions in Ho Chi Minh City, Vietnam. Purposive sampling was applied to guarantee that all respondents possessed a minimum of two years of direct experience preparing students for high-stakes graduation and professional language benchmarks.The sample exhibited a diverse range of instructional experience (from early-career lecturers to senior faculty members) and institutional profiles, with all participants holding postgraduate degrees (Master's or Ph.D. level). To protect participant confidentiality, all institutional and individual identifiers were fully anonymized through the use of alphanumeric codes and generic placeholders.

3.3. Instruments and data analysis

Quantitative data were gathered using a closed-ended questionnaire featuring a 5-point Likert scale (ranging from 1 = Strongly Disagree to 5 = Strongly Agree). The measurement items were adapted from established technology adoption instruments (Richter & Richter, 2024) and tailored to the operational realities of high-stakes IELTS teaching. The survey measured five core dimensions: pedagogic beliefs, technological knowledge, personal interest/usefulness, time constraints, and institutional culture. Quantitative analysis was executed using descriptive statistical software to compute means (M), standard deviations (SD), and standard errors. Qualitative insights were elicited via semi-structured, semi-formal interviews with a subset of ten participants (n=10) who volunteered during the survey phase. The interview protocol focused on participants' specific instructional adjustments, experiences with algorithmic discrepancies, and strategies for managing student dependence. All qualitative audio data were transcribed verbatim and analyzed using inductive thematic analysis (Braun & Clarke, 2021). This involved generating initial open codes, grouping codes into stable conceptual categories, and extracting overarching thematic dimensions to triangulate the quantitative findings.

4. Results

The mixed-methods analysis revealed that the pedagogical integration of AI-supported technologies in high-stakes IELTS preparation is not a linear adoption process. Instead, it is driven by a complex matrix of internal teacher cognitions *(beliefs, agency, and technological competence)* and external systemic constraints *(time, workload, and institutional policy).*

4.1. Quantitative Survey Findings

The quantitative data derived from the 50 surveyed EFL lecturers indicated varied levels of agreement across the major operational factors affecting technological adoption. The descriptive statistics for the five core dimensions are presented in Table 1.

Table 1

Factors influencing AI and Digital technology

Dimension	N	Mean	Std. Deviation	Std. Error Mean
Factor 1: Pedagogic beliefs & Alignment	50	3.84	0.62	.08
Factor 2: Technological knowledge & Competence	50	3.42	0.78	.11

Factor 3: Personal interest & Perceived usefulness	50	4.12	0.55	.07
Factor 4: Availability of time & Manageable workload	50	2.15	0.84	.12
Factor 5: Institutional culture & Policy support	50	2.89	0.91	.13

*Scores are based on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree).

As illustrated in Table 2, the highest mean score was recorded for Factor 3 (M = 4.12, SD = 0.55), indicating a highly positive internal orientation toward the technical capabilities of modern AI tools. Lecturers strongly perceive AI as a valuable asset for IELTS preparation. Conversely, the lowest score was observed for Factor 4 (M = 2.15, SD = 0.84). This indicates that severe professional time limits and heavy teaching schedules act as significant barriers to the exploration and practical integration of digital toolsets. Furthermore, Factor 5 scored moderately low (M = 2.89, SD = 0.91), validating the premise that AI adoption is currently occurring in a policy vacuum, forcing teachers to rely heavily on individual instructional agency (IA) rather than structured institutional guidance.

4.2. Qualitative interview themes

4.2.1. Efficiency gains vs. Algorithmic reliability

Lecturers widely acknowledged that AI applications significantly accelerated test-prep routines, particularly for generating topic-specific vocabulary and providing initial grammatical corrections. However, severe doubts persisted regarding the linguistic accuracy of generative AI feedback when applied to complex IELTS tasks. Participants noted that while AI excels at surface-level mechanics, it struggles with the rhetorical depth required for advanced band scores.

"AI is fantastic for quickly generating a list of collocations related to global supply chains for my logistics students. It saves me hours of preparation. But when I put a student's Task 2 essay into the software, the AI grades it a 7.5 based on vocabulary alone, completely ignoring the fact that the student's central argument lacks logical progression. It cannot read like a human examiner". (Participant 04)

4.2.2. Student dependence in pragmatic cohorts

A dominant theme was the rising anxiety over student over-dependence on generative AI. Students are often highly pragmatic, tech-savvy, and focused on efficiency. Consequently, they are highly susceptible to utilizing AI as an academic shortcut rather than a cognitive scaffold. This threatens the development of authentic communicative competence.

"My logistics students are brilliant at using technology to solve problems efficiently. The downside is that they treat IELTS writing as just another task to outsource. Instead of using ChatGPT to outline ideas for a manufacturing prompt, they ask it to write the whole essay, memorize it, and try to replicate it. They are submitting to the technology rather than learning from it, which destroys their independent language processing". (Participant 11)

4.2.3. The burden of Instructional Agency

The qualitative data strongly reflected the quantitative finding regarding low institutional support (M = 2.89). In the absence of clear departmental guidelines or systematic AI literacy training, the burden of regulating AI use falls entirely on the individual teacher's instructional agency.

"We are pressured to modernize our classrooms and ensure these high-quality cohorts hit their IELTS 6.5 benchmarks, but there is no official university policy on AI plagiarism or ethical use. I have to invent my own rules and boundaries for every writing assignment. It is exhausting to police this without institutional backing". (Participant 07)

5. Discussion

The empirical results of this study offer a nuanced examination of how tertiary EFL teachers reconcile the immense promises of AI with the rigid realities of high-stakes testing. The findings validate the theoretical premise that the pedagogical integration of AI is not determined by technological availability, but is deeply mediated by teacher cognition, instructional agency, and structural support.

5.1. The dissonance of PU and SC

The exceptionally high mean for *Personal Interest and Perceived Usefulness* (M = 4.12) demonstrates that Vietnamese EFL lecturers possess a strong progressive mindset toward educational technology. They clearly recognize AI's profound potential to revolutionize traditional, labor-intensive test preparation by providing real-time, low-stakes environments for lexical experimentation. However, this optimism is severely undercut by structural realities. The exceptionally low score for *Time Constraints* (M = 2.15) and the qualitative frustrations regarding workload reveal a critical operational bottleneck. Teachers recognize the value of AI, but lack the professional bandwidth required to rigorously test algorithmic reliability, adapt their lesson plans, and integrate these tools thoughtfully. This dissonance highlights that expecting educators to seamlessly absorb AI integration into their existing, heavy teaching schedules is an unsustainable institutional strategy.

5.2. Instructional Agency and the Challenges

The qualitative findings underscore the central role of Instructional Agency (IA) in navigating the risks of algorithmic dependence. The anxiety surrounding student over-reliance is particularly acute when instructing non-language majors, such as logistics students, who often prioritize task completion and operational efficiency over the slow, recursive process of language acquisition. In these high-quality tracks, where graduates must operate independently in unscripted, global professional landscapes, authentic communicative competence is non-negotiable.

When students use generative AI to synthesize entire responses, their apparent proficiency becomes a mere artifact of the software. To counteract this, teachers must exercise high instructional agency by deliberately setting pedagogical boundaries—such as restricting AI use to the brainstorming phase or requiring students to critically critique AI-generated essays against official IELTS rubrics. This agentic, selective integration proves that AI cannot replace the contextual, human expertise of an EFL educator; rather, it demands that the teacher evolves into a critical digital media mediator.

5.3. The Institutional support and Professional Development

The moderation effect of Institutional and Ethical Support (IES) is clearly visible in the teachers' narratives. The relatively low quantitative score for institutional culture (M = 2.89) correlates directly with the qualitative exhaustion teachers expressed regarding the "policy vacuum". When institutional governance is weak, the entire burden of defining ethical AI use and managing academic integrity falls on individual instructors. This approach creates inconsistencies across classrooms and increases teacher burnout.

To move beyond superficial technology adoption, universities must transition from passive observation to active structural support. This necessitates systematic, institutionalized PD that goes beyond basic software training. Effective PD must focus on AI literacy and digital pedagogy, teaching educators how to deconstruct algorithmic biases, align automated feedback with human IELTS rubrics, and design assessments that are resilient to generative AI shortcuts.

Furthermore, clear, university-wide ethical guidelines must be established to provide teachers with the institutional backing needed to confidently exercise their instructional agency.

5.4. Overreliance on AI writing

Identifying the threat of algorithmic dependence is insufficient without offering actionable pedagogical countermeasures. To mitigate the overreliance of pragmatic, tech-savvy logistics students, EFL lecturers must deploy proactive digital pedagogy solutions that reclaim instructional agency and enforce authentic cognitive engagement. *First,* teachers can adopt a *Flipped-AI-Framework*, which strictly confines generative AI tools to the pre-writing phase. While students may use AI to rapidly generate logistics-specific collocations or map structural outlines, the actual drafting must be conducted in-class without digital assistance, ensuring the cognitive effort of translating ideas into academic English remains entirely with the learner. *Second,* educators can implement *"Reverse Evaluation Tasks"* to foster critical AI literacy. By requiring students to generate an essay using AI and actively critique its flaws against official IELTS public band descriptors, teachers shift the student's role from a passive consumer to a critical evaluator. *Finally,* moving towards *"Process-Oriented Assessment and Verbal Defenses"* neutralizes the shortcut of automated text generation. Instructors should evaluate drafting iterations rather than just the final product and implement brief, randomized verbal defenses where students orally explain their lexical or grammatical choices. Together, these strategies act as powerful deterrents against submitting unedited, AI-generated work and foster genuine communicative competence.

6. Conclusion

This study investigated the core factors influencing how Vietnamese tertiary EFL lecturers perceive and adopt AI-supported and digital technologies within high-stakes IELTS preparation frameworks, specifically for logistics-majored student cohorts in high-quality academic tracks. The findings demonstrate that technology integration in this specialized context is a highly complex, non-linear process mediated by teacher cognition, instructional agency, and structural constraints.

While lecturers exhibit strong personal interest and clearly recognize the operational utility of AI in accelerating feedback loops and enhancing vocabulary generation, their pedagogical adoption is heavily hindered by severe time constraints and a pervasive institutional policy vacuum. Furthermore, the unique pragmatic characteristics of logistics students amplify profound concerns regarding automated reliability and algorithmic dependence. When students utilize generative AI as an academic shortcut rather than a cognitive scaffold, they bypass the recursive learning processes essential for authentic language acquisition, risking misalignment with strict human-evaluated IELTS rubrics.

To ensure that AI functions as a legitimate driver of deep language learning, higher education institutions must urgently transition away from ad-hoc technological usage. It is imperative that universities implement systematic, institutionalized professional development (PD) focused on digital pedagogy and AI literacy, empowering teachers to critically evaluate algorithmic outputs. Concurrently, establishing transparent, university-wide ethical guidelines (IES) is essential to relieve the regulatory burden on individual educators. Only through deliberate, agentic integration, supported by robust institutional frameworks, can EFL practitioners maximize the transformative potential of AI while preserving the academic rigor required for sustainable, high-level English communicative competence in global professional landscapes.

References

An, F., Yu, J., & Xi, L. (2022). Relationship between perceived teacher support and learning engagement among adolescents: Mediation role of technology acceptance and learning motivation. *Frontiers in Psychology*, *13*, 992464. https://doi.org/10.3389/fpsyg.2022.992464

Azis, M. A., & Numan, M. (2026). Teachers' beliefs and decisions regarding artificial intelligence use. *Duties: Education and Humanities International Journal*, *2*(1), 69-91. https://doi.org/10.70152/duties.v2i1.323

Bakshi, P. (2021). Are teachers boundedly rational? How beliefs and context affect decision-making. *Leadership and Policy in Schools*, *22*(1), 261-282. https://doi.org/10.1080/15700763.2021.1950773

Balzano, M., Marzi, G., & Turzo, T. (2024). SMEs and institutional theory: Major inroads and opportunities ahead. *Management Decision*, *63*(13), 1-27. https://doi.org/10.1108/MD-05-2023-0734

Bower, M., Torrington, J., Lai, J. W. M., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, *29*, 15403-15439. https://doi.org/10.1007/s10639-023-12405-0

Chiu, T. K. F., Ahmad, Z., & Çoban, M. (2025). Development and validation of teacher artificial intelligence (AI) competence self-efficacy (TAICS) scale. *Education and Information Technologies*, *30*, 6667-6685. https://doi.org/10.1007/s10639-024-13094-z

Cong-Lem, N., Tran, T. N., & Nguyen, T. T. (2024). Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers. *Teaching English With Technology*, *2024*(1). https://doi.org/10.56297/fsyb3031/mxnb7567

Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, *10*, 12. https://doi.org/10.1186/s40561-023-00231-3

Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*, *20*(1), 3. https://doi.org/10.1007/s40979-024-00153-8

Latif, M. W., & Wasim, A. (2022). Teacher beliefs, personal theories and conceptions of assessment literacy - A tertiary EFL perspective. *Language Testing in Asia*, *12*(1), 11. https://doi.org/10.1186/s40468-022-00158-5

Lei, W., & Lei, Z. (2025). Exploring EFL teachers' beliefs and practices of formative assessment in Chinese context. *PLOS One*, *20*(9), e0333678. https://doi.org/10.1371/journal.pone.0333678

Li, J., & Huang, Q. (2025). Navigating ethical dilemmas in AI-enhanced education. *International Journal of Knowledge Management*, *21*(1). https://doi.org/10.4018/IJKM.382384

Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English language learners: A systematic review. *Information*, *16*, 519. https://doi.org/10.3390/info16070519

Mohammad Nezhad, P., & Stolz, S. A. (2025). Unveiling teachers' professional agency and decision-making in professional learning: The illusion of choice. *Professional Development in Education*, *51*, 1067-1087. https://doi.org/10.1080/19415257.2024.2349058

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2025). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*, *30*(3), 3343-3387. https://doi.org/10.1007/s10639-024-12946-y

Wu M, Yao J and Al-Hashimy HNH (2026). English teachers' perceptions of AI-supported writing instruction: Implications for professional development. *Frontiers in Psychology*, *17*, 1806618. https://doi.org/10.3389/fpsyg.2026.1806618

Zhang, Y., & Li, C. (2025). Leveraging pedagogical knowledge for effective technology integration in education: A research-based approach. *International Journal of Educational Research*, *130*, 102553. https://doi.org/10.1016/j.ijer.2025.102553

Bionote

Thai Ho-Van is a lecturer at the Faculty of Foreign Languages, College of Foreign Economic Relations (COFER) in Ho Chi Minh City, Vietnam. Holding a TESOL Master's degree, PhD candidate at Open University of Ho Chi Minh City, his primary research explores innovative educational methodologies, including TBL, CLIL, and EMI. His current academic focus investigates the critical integration of AI-driven tools for students, aiming to enhance English competence in higher education.

#92 BEYOND ACCURACY: PRAGMATIC FORCE IN AI-MEDIATED TRANSLATING AND INTERPRETING TRAINING

Dương Thị Hoàng Oanh, Nguyễn Xuân Đạt

Abstract

In the context of increasing attention to assessing communicative competence and real-world language use, this study explores the role of pragmatic force in AI-mediated translating and interpreting (T&I) training and assessment in English as a Second Language/ English as a Second Language (ESL/EFL) contexts, with particular relevance to language education in Vietnam. While recent advances in Artificial Intelligence (AI) have improved the accuracy and efficiency of T&I, they remain limited in capturing essential aspects of communication such as speaker/writer intention, implicature, stance, and interactional meaning. This paper therefore reconceptualises T&I assessment beyond linguistic accuracy by positioning pragmatic force as a central component of meaningful communication in technology-supported environments. Drawing on current literature and classroom-informed practices, the study examines how human–machine collaboration is reshaping T&I into human-in-the-loop models, where translators and interpreters actively monitor, adjust, and intervene to ensure pragmatic appropriateness. It also proposes practical pedagogical strategies for ESL/EFL educators, including the use of authentic, simulation-based tasks and AI-supported activities that reflect real-life communicative demands. The findings emphasise that communicative competence in T&I involves not only accurate language transfer but also the ability to interpret and convey meaning appropriately in context, especially in AI-mediated settings. The study suggests that T&I education in ESL/EFL classrooms should prioritise pragmatic awareness, critical use of AI tools, and flexible decision-making skills. In conclusion, placing pragmatic force at the heart of training and assessment can better equip students to communicate effectively and confidently in today's increasingly technology-rich professional environments.

Keywords: Pragmatic Force; Communicative Competence; AI-Mediated Tools; ESL/EFL; Translator/Interpreter Training

1. Introduction

The ability to use language appropriately in context is widely recognized as a fundamental component of communicative competence. This insight remains highly relevant in today's rapidly evolving language landscape, where advances in Artificial Intelligence (AI) are transforming the practice and training of translating and interpreting (T&I). AI-powered tools such as neural machine translation, speech recognition systems, computer-assisted translation technologies, and large language models (LLMs) have significantly enhanced the speed, accessibility, and efficiency of multilingual communication. These technologies are increasingly integrated into professional workflows and educational settings, offering valuable support for language learners, translators, and interpreters alike.

Despite these technological advances, concerns remain regarding the ability of AI systems to capture the deeper dimensions of human communication. While AI can often generate linguistically fluent and seemingly accurate translations and interpretations, it frequently struggles with aspects of meaning that depend heavily on context, culture,

interpersonal relationships, and communicative intent. Nuances such as implicature, politeness, stance, irony, persuasion, and speaker intention may be overlooked or inadequately conveyed. Consequently, a translation or interpretation that appears accurate at the lexical or grammatical level may nevertheless fail to achieve its intended communicative effect. This limitation foregrounds **pragmatic force**, the intended meaning and communicative impact of an utterance, as a critical yet often underrepresented dimension of T&I performance.

Although an increasing body of research has examined AI-assisted translation, machine translation quality, translator competence, and human, AI collaboration (Bowker, 2023; O'Hagan, 2020), comparatively little attention has been paid to how **pragmatic force should be conceptualized, taught, and assessed in AI-mediated T&I environments**. Existing studies tend to focus on linguistic accuracy, productivity, technological efficiency, or post-editing performance, while the communicative consequences of pragmatic misalignment remain underexplored. Similarly, assessment frameworks in translation and interpreting continue to privilege semantic equivalence and linguistic correctness, often providing limited guidance on how pragmatic competence can be evaluated in contexts where AI increasingly participates in the communication process. **This represents an important gap in literature, particularly within ESL/EFL educational settings where communicative competence is a central learning outcome.**

The problem addressed in this paper is that current T&I training and assessment practices do not adequately account for pragmatic force in AI-mediated communication. As AI technologies become embedded in educational and professional contexts, learners may produce or rely on linguistically accurate outputs that fail to convey intended meanings, interpersonal relationships, or culturally appropriate messages. Without explicit attention to pragmatic competence, students risk developing technological proficiency without acquiring the communicative judgment necessary to evaluate, adapt, and improve AI-generated translations and interpretations. This challenge is particularly significant in ESL/EFL contexts, where learners must simultaneously develop language proficiency, intercultural awareness, and digital literacy.

The issue is especially relevant in Vietnam. Recent educational reforms have emphasized communicative competence and international engagement, while national initiatives promoting English language proficiency have increased the importance of effective multilingual communication. At the same time, AI technologies such as ChatGPT, Google Translate, Gemini, and other language-support tools are being adopted rapidly by students and educators. As graduates prepare for increasingly technology-rich workplaces, they require not only linguistic competence but also the ability to critically evaluate AI-generated outputs, recognize pragmatic mismatches, and make informed decisions to ensure communicative effectiveness across languages and cultures.

This paper is conceptual in nature and focuses specifically on translation and interpreting training and assessment in higher-education ESL/EFL contexts, with particular reference to Vietnam. It does not attempt to evaluate the technical performance of specific AI systems, compare machine translation engines, or provide empirical measurement of learner outcomes. Rather, it examines how insights from pragmatics, communicative competence, translation studies, and AI-assisted language learning can be integrated to reconceptualize T&I education in the age of AI. Accordingly, the discussion is limited to educational and pedagogical implications for translator and interpreter training, rather than professional industry practices beyond the educational context.

Against this backdrop, the paper reconceptualizes T&I training and assessment by positioning pragmatic force at the heart of communicative competence in AI-mediated environments. It examines how human–machine collaboration is reshaping T&I into human-in-the-loop models, where translators and interpreters actively monitor, revise, and intervene to ensure pragmatic appropriateness. The paper further proposes pedagogical strategies to promote pragmatic awareness, critical AI literacy, and context-sensitive decision-making. By moving beyond accuracy as the sole benchmark of quality, the study argues for a more comprehensive understanding of T&I competence that better prepares learners to communicate effectively and professionally in an increasingly AI-driven world.

2. From Linguistic Accuracy to Communicative Competence in Translating and Interpreting

2.1 Theoretical Framework: Understanding Pragmatic Force in AI-Mediated T&I

2.1.1 Speech Act Theory: Meaning Beyond Words

This study is primarily grounded in **Speech Act Theory** (Austin, 1962; Searle, 1969), which views language as action rather than merely a vehicle for transmitting information. Austin (1962) argues that communication operates on three interconnected levels: the **locutionary act** (what is said), the **illocutionary act** (what is intended), and the **perlocutionary act** (the effect on the audience). From this perspective, successful translation and interpreting involve more than transferring linguistic content; they require preserving the intended communicative force of the original message.

Speech Act Theory provides a particularly useful lens for examining AI-mediated T&I because many translation errors arise not from lexical inaccuracies but from failures to convey illocutionary force. A machine-generated translation may accurately reproduce propositional content while misrepresenting requests, warnings, refusals, apologies, or persuasive intentions. Consequently, this theory provides the primary conceptual foundation for the notion of pragmatic force adopted in this study.

2.1.2 Systemic Functional Linguistics: Language as Social Meaning

To complement Speech Act Theory, the study draws upon Systemic Functional Linguistics (SFL) (Halliday, 1978; Halliday & Matthiessen, 2014), which conceptualizes language as a resource for meaning-making within social contexts. SFL proposes three metafunctions of language:

- Ideational metafunction: representing experience and content;
- Interpersonal metafunction: establishing relationships, attitudes, and social positioning;
- Textual metafunction: organizing messages into coherent discourse.

Among these, the interpersonal metafunction is particularly relevant to pragmatic force because it explains how speakers express politeness, authority, solidarity, evaluation, and stance. In T&I, successful communication requires preserving not only ideational meaning but also interpersonal meanings that shape audience interpretation. SFL therefore provides a framework for analyzing how pragmatic force is realized linguistically and socially.

2.1.3 Communicative Competence: Appropriateness in Context

The third theoretical pillar is Communicative Competence Theory (Hymes, 1972; Canale & Swain, 1980; Bachman & Palmer, 2010). This perspective extends language ability

beyond grammatical knowledge to include sociolinguistic, pragmatic, and strategic competences.

Within this framework, pragmatic competence refers to the ability to interpret and use language appropriately according to context, audience, and communicative purpose. For translators and interpreters, communicative competence involves not only understanding linguistic meanings but also recognizing intentions, managing intercultural differences, and making context-sensitive decisions. This perspective is particularly relevant in AI-mediated environments where human professionals increasingly serve as evaluators and mediators of machine-generated communication.

2.1.4 Conceptual Model Guiding the Study

Drawing upon these three theoretical perspectives, this paper conceptualizes successful translation and interpreting as the preservation of pragmatic force across languages and cultures.

Specifically:

Theory	Contribution to This Study
Speech Act Theory (Austin, 1962; Searle, 1969)	Explains speaker intention and communicative force
Systemic Functional Linguistics (Halliday, 1978)	Explains how interpersonal meanings are realized in discourse
Communicative Competence Theory (Hymes, 1972; Canale & Swain, 1980)	Explains appropriateness and effectiveness in context

Together, these theories underpin the central argument that successful T&I should be assessed not only according to semantic accuracy but also according to the extent to which the translated or interpreted message preserves the communicative intentions, interpersonal meanings, and audience effects of the original message.

3. Understanding Pragmatic Force in Translation and Interpreting (T&I)

3.1 Analytical Approach: Literature-Informed Thematic Analysis of Classroom Practices

Although this paper is primarily conceptual, the discussion is informed by both contemporary literature and recurring observations from translation and interpreting classrooms in Vietnamese higher-education ESL/EFL contexts. To enhance analytical transparency, examples and pedagogical implications presented throughout the paper were interpreted using a literature-informed thematic analysis approach (Braun & Clarke, 2006, 2022). Rather than testing hypotheses, thematic analysis was used to identify recurring communicative challenges observed in student translations, interpretations, classroom discussions, reflective tasks, and AI-assisted learning activities.

Through iterative review of these classroom-informed observations alongside relevant literature, several recurring themes emerged, including: (1) difficulties in recognizing speaker intention; (2) challenges in interpreting conversational implicatures; (3) inappropriate rendering of politeness and interpersonal stance; (4) misinterpretation of culturally sensitive references; and (5) overreliance on AI-generated outputs without sufficient pragmatic evaluation. These themes provide a practical lens through which the concept of pragmatic force

is examined and help connect theoretical discussions to pedagogical realities in AI-mediated T&I education.

Thematic analysis is particularly appropriate because pragmatic force often manifests through patterns of communicative misunderstanding rather than isolated linguistic errors. Accordingly, the following discussion integrates theoretical perspectives from pragmatics and translation studies with representative classroom-informed examples to illustrate how pragmatic force influences translation and interpreting quality.

3.2 Defining Pragmatic Force

While linguistic accuracy focuses on the transfer of lexical, grammatical, and semantic content, successful communication depends equally on the ability to convey what speakers or writers intend to achieve through language. This communicative effect is often referred to as pragmatic force, the intended meaning, interpersonal stance, and social action associated with an utterance within a particular context.

Pragmatic force encompasses a range of communicative dimensions, including speaker intention, implicature, politeness, evaluation, emotional stance, and audience orientation. These dimensions were also among the most frequently recurring themes identified in classroom-informed observations. In particular, students often demonstrated strong lexical comprehension while encountering difficulties in recognizing implied meanings, interpersonal positioning, and culturally embedded communicative intentions. Such recurring patterns reinforce the view that pragmatic competence extends beyond linguistic knowledge and requires sensitivity to contextual and sociocultural factors.

Pragmatic competence is essential for effective communication in any language pair (Darong 2024). For example, to translate the headline “Inside the secret mission to fly Taiwanese’s president to Africa” into Vietnamese, ChatGPT may instinctively translate “president” as “tổng thống”; such a literal translation risks neglecting pragmatic considerations and broader sociopolitical sensitivities associated with cross-strait and regional diplomatic relations involving Vietnam, China, and Taiwan: the official translation of the title into Vietnamese is “nhà lãnh đạo Đài Loan” (the leader of Taiwan). This demonstrates that pragmatic force in translation extends beyond lexical equivalence to preserving communicative intentions, contextual meanings, and appropriate social implications (Dayter et al. 2023). Therefore, meaning is shaped not only by linguistic forms, but also by contextual factors, social relationships (Leech 1983), and political viewpoints, with the meaning of an utterance extending beyond its literal semantic content to include how it is interpreted within a specific communicative situation. Similar challenges emerged in classroom activities involving AI-generated translations, where students frequently accepted linguistically fluent outputs without critically examining their political, cultural, or interpersonal implications. Thematic analysis of classroom discussions suggested that such instances were less related to vocabulary deficiencies than to difficulties in evaluating context-sensitive meanings and communicative consequences.

As House (2015) argues, translation quality cannot be assessed solely in terms of linguistic equivalence because successful communication also requires attention to functional and pragmatic appropriateness. A translation that accurately reproduces lexical meaning may nevertheless fail if it does not convey the intended interpersonal relationship, level of formality, or communicative purpose. For translators and interpreters, pragmatic force is particularly important because they serve as mediators between languages, cultures, and communicative communities. Their task is not only to transfer information but also to recreate the intended

effect of the source message for a new audience. Consequently, effective T&I requires sensitivity to contextual factors that shape meaning beyond words.

One way to assess whether a T&I successfully achieves pragmatic force is by observing the audience's response rather than checking literal accuracy (Matsushita 2026). If the target audience reacts similarly to the original audience, then the T&I can be considered successful.

It is noteworthy that a translated text or interpreted speech can be regarded as pragmatically successful when it generates an effect on the target audience that is functionally equivalent to, or closely resembles, the effect of the original on its intended audience. This means that successful translation and interpreting involve more than the accurate transfer of words and meanings; they require the preservation of communicative intent, interpersonal relationships, cultural nuances, and the overall pragmatic force of the source message. Achieving this goal requires translators and interpreters to make informed decisions regarding context, culture, audience expectations, and communicative purpose.

Pragmatic force remains a central criterion for successful communication. While AI technologies may assist with rapid linguistic processing and terminology retrieval, interpreters must still rely on contextual adaptation, clarification, and ethical judgement to preserve pragmatic fidelity. Effective interpreting therefore depends on balancing linguistic accuracy with cultural and interpersonal awareness. The future of interpreting likely depends on a hybrid model in which AI supports efficiency, while human interpreters safeguard communicative trust and socio-cultural meaning (Kokorelias et al. 2026).

3.3 Pragmatic Failure in Translation and Interpreting

The importance of pragmatic force is rooted in speech act theory. Austin (1962) distinguished among three dimensions of meaning. The locutionary act refers to the literal linguistic content of an utterance, the words and sentences that are spoken. The illocutionary act refers to the speaker's intended function, such as requesting, promising, apologizing, warning, or inviting. The perlocutionary act concerns the effect the utterance has on the listener, such as persuading, reassuring, offending, or motivating. Effective T&I requires the ability to infer and reproduce illocutionary force rather than focusing exclusively on literal content. Failure to recognize pragmatic force can result in what Thomas (1983) termed pragmatic failure, defined as "a failure to understand what is meant by what is said" (p. 91). Unlike grammatical errors, pragmatic failures may occur even when every word is translated accurately. The problem lies not in linguistic form but in the interpretation of communicative intention, and in the effect to the target language audience.

Thomas (1983) distinguished between pragmalinguistic failure, which involves incorrect interpretation of linguistic strategies, and sociopragmatic failure, which arises from differing cultural expectations and social norms. This distinction was reflected in recurring classroom observations: pragmalinguistic difficulties frequently appeared when students interpreted indirect requests, suggestions, or hedging expressions literally, while sociopragmatic difficulties were more commonly observed in situations involving differing expectations regarding politeness, hierarchy, relational pronouns, and culturally appropriate forms of address. Such patterns emerged repeatedly across interpreting simulations and AI-assisted translation tasks, suggesting that pragmatic failure remains a persistent challenge even when linguistic accuracy is relatively high. Both types are highly relevant to T&I. For example, an English speaker may say "Let's have coffee some time!" as a general friendly phrase with no concrete expectations, whereas other cultures may interpret that as an opening to make concrete plans. Similar issues arise in English-Vietnamese translation and interpreting. For

example, Vietnamese relational pronouns such as anh, chị, cô, chú, or em carry interpersonal meanings related to age, status, and social relationships that are often absent in English. Students using AI-generated translations frequently reported difficulties determining the most appropriate forms of address because machine-generated outputs tended to prioritize lexical equivalence rather than relational positioning. Thematic analysis suggested that these choices often influenced the perceived politeness and interpersonal tone of the translated message.

The importance of pragmatic force becomes even more evident in AI-mediated translation and interpreting. While AI systems have become increasingly proficient at generating fluent and semantically accurate output, they frequently struggle with irony, humour, indirect requests, politeness conventions, cultural references, and implied meanings. As a result, machine-generated translations may appear accurate while failing to achieve the intended communicative effect. For example, the English expression “That's just great” may convey genuine approval or, depending on context and intonation, sarcasm and frustration. An AI system may translate the phrase literally into another language as positive praise while missing the speaker's intended criticism.

Human translators and interpreters therefore remain indispensable in identifying pragmatic nuances, resolving ambiguities, and ensuring that translated messages fulfill their intended purposes. Understanding pragmatic force is thus essential not only for translation quality but also for the effective integration of AI into contemporary T&I practice and education.

4. AI in Translating and Interpreting: Opportunities and Pragmatic Limitations

4.1 Advances in AI Tools for Translating and Interpreting

The rapid development of artificial intelligence (AI) has transformed the field of T&I. Early rule-based and statistical machine translation systems have evolved into sophisticated neural architectures capable of generating fluent multilingual output. Today, translators and interpreters routinely use machine translation (MT), computer-assisted translation (CAT) tools, automatic speech recognition (ASR), terminology management systems, and LLMs such as ChatGPT, Gemini, and Claude.

Neural Machine Translation (NMT) has significantly improved translation quality by processing entire sentences and broader contextual information rather than translating word by word. CAT tools further enhance productivity through translation memories, terminology databases, and quality assurance functions, while advances in speech recognition and synthesis have expanded possibilities for real-time multilingual communication.

According to O’Hagan (2020), these technologies have become integral to contemporary language services, reshaping professional workflows and translator competences. Rather than replacing human practitioners, AI increasingly operates within collaborative human-machine environments.

4.2 Strengths of AI in Translation and Interpreting

AI technologies offer several advantages for translation and interpreting. Their most significant strengths are speed, consistency, and scalability. Tasks that once required substantial human effort can now be completed in seconds, enabling translators to process larger volumes of text and meet demanding deadlines.

AI systems also improve terminological consistency through translation memories and terminology management tools, which are particularly valuable in legal, medical, scientific,

and business translation. In addition, contemporary NMT systems can produce highly fluent and grammatically accurate output for routine and standardized content. As Koehn and Knowles (2017) note, recent advances in NMT have generated translations that appear increasingly natural and human-like.

4.3 Limitations in Conveying Pragmatic Force

Despite their impressive capabilities, AI systems face significant limitations when dealing with pragmatic meaning that depends not on linguistic form alone but on context, intention, culture, and interpersonal relationship. These limitations are worth examining systematically, as they define precisely where human expertise remains indispensable.

The most frequently cited problem is the gap between fluency and understanding. As Koehn and Knowles (2017) caution, neural MT systems are "remarkably fluent but not necessarily accurate in meaning" (p. 28), and fluent output can create a dangerous illusion of correctness. This becomes particularly acute when the intended meaning is indirect, embedded in cultural knowledge, or deliberately ambiguous. A striking example from Vietnamese literature illustrates the point: the phrase "Mày râu nhẵn nhụi, áo quần bảnh bao" would be rendered by most MT systems as something like "clean-shaven and smartly dressed" – a positive description. In context, however, it is a reference to the villain Sở Khanh in The Tale of Kiều, and functions as a culturally loaded warning about deceptive appearances. No amount of linguistic fluency recovers what the allusion is doing.

Politeness, hierarchy, and interpersonal stance present a related but distinct challenge. Languages encode social relationships very differently, and machine translation routinely flattens these distinctions. Vietnamese pronouns are a well-documented case: a speaker's choice among anh, chị, em, bác, cô, and dozens of other forms encodes age, gender, relative status, intimacy, and institutional context simultaneously. Rendered uniformly as "I" and "you" in English, the relational texture of the original (who is deferring to whom, who is claiming authority, what kind of relationship is being performed) disappears entirely. The propositional content survives; the social meaning does not.

Conversational implicature poses a further difficulty. Human communicators routinely infer meaning that is not explicitly stated, drawing on shared knowledge, situational context, and cultural expectations. AI systems, processing language through statistical pattern-matching rather than communicative understanding, frequently miss what a speaker intends to imply rather than say. Bender and Koller (2020) put the theoretical point sharply: language models manipulate linguistic form without possessing the experiential and cultural knowledge humans draw on when constructing meaning (p. 518). A model can recognize the words of an indirect refusal, an ironic compliment, or a face-saving hedge and still render them as if they were literal.

These limitations are compounded in cross-cultural T&I, where communication regularly requires adaptation rather than transfer. Humour, political allusion, institutional register, and culturally specific norms of appropriate speech may have no direct equivalent in the target language or culture. Consider diplomatic interpreting: a speaker's careful use of hedged, non-committal language may be a deliberate political act, signalling openness without commitment, or maintaining plausible deniability. A literal rendering that preserves the hedges may read as evasive in the target culture; one that smooths them out may inadvertently commit the speaker to a stronger position than intended. Human interpreters navigate this in real time; current AI systems lack the situational awareness to do so reliably.

For these reasons, scholars increasingly advocate a human-in-the-loop model of AI-assisted T&I – one in which AI functions as a powerful support tool whose output requires human evaluation, revision, and contextual judgement. The question is not whether AI can replace translators and interpreters, but how human expertise can be deployed most effectively alongside AI: identifying pragmatic failures, resolving ambiguities, and ensuring that translated messages achieve their intended communicative effect rather than merely their linguistic form.

5. Human-in-the-Loop: Reframing Translation and Interpreting as Collaborative Practice

5.1 The Concept of Human-in-the-Loop Translation and Interpreting

The growing integration of AI into T&I has prompted a fundamental reconsideration of the relationship between human expertise and technological assistance. Rather than viewing AI as a replacement for translators and interpreters, contemporary scholarship increasingly conceptualizes T&I as a collaborative process in which humans and machines contribute complementary strengths, which can be described as human-in-the-loop translation and interpreting.

In this model, AI systems perform tasks such as generating draft translations, recognizing speech, retrieving terminology, or suggesting linguistic alternatives, while human translators and interpreters remain actively involved in evaluating, revising, and validating the output. The role of the human participant extends beyond error correction to include critical decision-making regarding meaning, context, audience expectations, and communicative effectiveness.

One of the most visible manifestations of this model is post-editing, in which machine-generated translations are reviewed and modified by human translators. Although post-editing is often associated with improving linguistic accuracy, it increasingly involves addressing deeper communicative issues that AI systems may fail to recognize. According to Vieira (2020), "post-editing requires not only linguistic but also pragmatic judgment," highlighting the need for translators to assess whether machine-generated output successfully conveys the intended communicative force. Post-editors must therefore evaluate not only what the text says but also what it means and how it is likely to be interpreted by target audiences.

This collaborative perspective reflects broader developments in translation technology. As AI systems become more capable and self-learning, the focus shifts from performing T&I manually to managing, supervising, and optimizing interactions between human expertise and machine assistance. Consequently, the value of translators and interpreters increasingly lies in areas where human cognition and social understanding remain indispensable.

5.2 The Translator and Interpreter as Decision-Maker

The human-in-the-loop paradigm redefines translators and interpreters not as passive users of technology but as active decision-makers. While AI can generate linguistic output rapidly and efficiently, it lacks the capacity to assume responsibility for communicative outcomes. Human practitioners remain accountable for ensuring that translated messages are accurate, appropriate, ethical, and effective.

O'Hagan (2020) captures this transformation succinctly by observing that "human translators are not replaced but repositioned." Rather than competing with machines, translators and interpreters assume new roles as evaluators, mediators, quality managers, and communication specialists. Their expertise becomes particularly important in situations involving ambiguity, cultural sensitivity, interpersonal relationships, and high-stakes communication.

This repositioning also highlights the importance of professional agency. Translators and interpreters routinely make decisions regarding lexical choice, register, tone, politeness, audience adaptation, and cultural references. Such decisions require contextual awareness and communicative judgment that cannot be reduced to algorithmic procedures. In AI-mediated environments, professionals must determine when machine-generated output is acceptable, when revision is necessary, and when entirely different solutions are required.

Ethical responsibility constitutes another crucial dimension of human agency. AI systems may inadvertently generate biased, misleading, culturally inappropriate, or contextually insensitive T&I. Human oversight is therefore essential to identify potential risks and ensure responsible communication. Translators and interpreters must evaluate not only linguistic quality but also the broader social, ethical and political implications of their choices. This responsibility becomes particularly significant in domains such as healthcare, law, diplomacy, education, and public communication, where misunderstandings may have serious consequences.

As Pym (2013) argues, the emergence of machine translation requires a reconceptualization of translator competence. Rather than focusing exclusively on language transfer, future professionals need skills related to evaluation, revision, technological literacy, and intercultural mediation. In this sense, technology increases rather than diminishes the importance of human expertise.

5.3 Pragmatic Mediation as a Distinctively Human Strength

Among the many competencies required in AI-assisted T&I, pragmatic mediation remains one of the most significant areas of human contribution. As discussed in previous sections, pragmatic force encompasses speaker intention, implicature, politeness, stance, and other context-dependent aspects of meaning. These dimensions often require interpretation that extends beyond the linguistic surface of a text.

Human translators and interpreters possess the ability to draw upon cultural knowledge, social experience, contextual awareness, and inferential reasoning when interpreting communicative intent. They can recognize irony, indirectness, humour, emotional undertones, and interpersonal dynamics that may not be explicitly encoded in language. This capacity enables them to adapt messages in ways that preserve communicative purpose across languages and cultures. In contrast, AI systems primarily rely on statistical patterns learned from large datasets. Although they can generate fluent and often impressive translations, they frequently struggle with pragmatic nuances that require an understanding of human intentions and social relationships. Consequently, pragmatic mediation remains a domain in which human expertise offers substantial added value.

From a communicative perspective, translators and interpreters function not merely as language converters but as meaning negotiators. Their role involves ensuring that target audiences understand messages in ways that approximate the interpretations intended for source-language audiences. This process often requires adapting tone, adjusting politeness levels, clarifying implied meanings, and making culturally and politically appropriate choices that extend beyond literal translation.

The human-in-the-loop model therefore underscores a crucial insight: the future of T&I is not a contest between humans and machines but a collaboration that combines technological efficiency with human communicative intelligence. While AI contributes speed, consistency, and linguistic support, human translators and interpreters provide the pragmatic judgment, ethical responsibility, and intercultural competence necessary for effective communication. As

AI becomes increasingly integrated into professional practice, these uniquely human capacities are likely to become even more central to T&I education, assessment, and professional identity.

If pragmatic judgment is increasingly recognized as a core human contribution in AI-mediated T&I, then training and assessment practices must evolve accordingly. The next section therefore explores how pragmatic force can be integrated into T&I pedagogy and assessment through authentic, AI-supported learning activities that reflect real-world communicative demands.

6. Pedagogical Implications for ESL/EFL Translation and Interpreting Training

The growing integration of AI into T&I requires a fundamental reconsideration of how future translators and interpreters are trained and assessed. If communicative competence in AI-mediated environments extends beyond linguistic accuracy to include pragmatic judgment, intercultural awareness, and critical engagement with technology, then educational practices must evolve accordingly. This need is particularly pressing in ESL/EFL contexts such as Vietnam, where language education is increasingly influenced by AI tools while in many institutions still remaining shaped by examination-oriented traditions and form-focused instruction. To prepare students for contemporary professional realities, T&I training should incorporate pragmatic awareness, AI literacy, authentic communicative tasks, and assessment frameworks that recognize pragmatic force as a key component of translation quality.

6.1 Integrating Pragmatic Awareness into T&I Training

A central implication of the present discussion is the need to place pragmatic competence at the heart of T&I education. While students often receive extensive training in grammar, vocabulary, and translation techniques, they may have limited opportunities to develop awareness of how meaning is shaped by context, intention, and social interaction.

Pragmatic instruction can focus on communicative features that frequently create difficulties in cross-cultural communication, including speech acts, implicature, politeness strategies, stance, register, hedging, and interpersonal positioning. Students should be encouraged to analyse not only what speakers say but also what they intend to accomplish through language. Activities may involve comparing literal meanings with intended meanings, identifying implied messages, examining politeness strategies across cultures, and evaluating how different linguistic choices influence audience interpretation.

For example, learners can compare multiple translations of the same utterance and discuss how each version affects interpersonal relationships, levels of formality, or communicative impact. Interpreting students may analyse authentic recordings to identify indirect requests, persuasive language, culturally or politically embedded meanings that require adaptation during interpretation. Such activities help students recognize that successful T&I involves more than transferring information; it requires understanding how language functions within specific social, cultural and political contexts. Developing this awareness is particularly important in AI-assisted environments, where machine-generated output may appear accurate while failing to convey intended pragmatic meanings.

6.2 AI-Mediated Classroom Practices

Rather than prohibiting AI tools, educators should incorporate them strategically into teaching and learning activities. As AI technologies become increasingly embedded in professional practice, students need opportunities to develop the skills required to use these tools critically and responsibly.

AI platforms such as ChatGPT, Google Translate, DeepL, and speech-recognition systems can serve as valuable pedagogical resources when used appropriately. However, the goal should not be to encourage dependence on technology but to foster critical evaluation and informed decision-making. As Chapelle (2003) reminds us, "Technology should support, not replace, pedagogy." The educational value of AI therefore depends on how it is integrated into meaningful learning experiences.

In our classes, one effective approach involves asking students to compare human and AI-generated translations. The students can identify strengths and weaknesses, evaluate pragmatic appropriateness, and revise machine-generated outputs to better suit specific communicative contexts. Such activities help students develop awareness of common AI limitations related to tone, politeness, audience adaptation, and cultural nuance.

Another possibility is the use of AI-generated dialogue simulations. Our students interact with AI systems acting as clients, conference participants, healthcare professionals, or business partners, thereby practicing translation and interpreting skills in realistic communicative scenarios. These activities allow the students to experiment with different strategies while receiving immediate feedback and exposure to diverse communicative situations.

In Vietnam, where opportunities for authentic interaction with English-speaking communities outside the classes may be limited, AI tools can provide valuable supplementary exposure to varied discourse types, registers, and communicative contexts. However, we would need to emphasize that AI functions as a scaffold for learning rather than a shortcut to performance. Students must learn to question, verify, and improve AI-generated output rather than accept it uncritically.

6.3 Authentic and Simulation-Based Learning Tasks

The development of communicative competence requires learning experiences that resemble real-world professional practice. Many traditional T&I exercises often focus on isolated texts or sentence-level accuracy, providing limited opportunities for students to engage with the contextual and interpersonal dimensions of communication.

According to Gulikers et al. (2004), "authentic tasks engage learners in real-world language use" (p. 69). Such tasks require students to apply knowledge and skills in contexts that closely resemble those encountered outside the classroom.

In T&I education, authenticity can be achieved through activities such as client-based translation briefs, multilingual business correspondence, public service translation projects, conference interpreting simulations, community interpreting role plays, and media translation assignments. Students can be asked to consider audience needs, communicative objectives, cultural expectations, and ethical considerations alongside linguistic accuracy. Interpreting simulations are particularly valuable because they require learners to process language in real time while responding to interpersonal dynamics and contextual cues. For example, a community interpreting role play involving a healthcare consultation may require students to manage politeness, cultural expectations, and sensitive information while ensuring accurate communication between participants.

AI technologies can further enrich these experience by generating realistic scenarios, diverse speaker profiles, and context-specific materials. When combined with human reflection and feedback, simulation-based activities help students develop the adaptive and context-sensitive decision-making skills required in professional practice.

6.4 Redesigning Assessment Beyond Accuracy

Perhaps the most significant pedagogical implication concerns assessment. If pragmatic force is central to effective communication, then T&I assessment would need to move beyond traditional measures of linguistic accuracy and incorporate criteria that evaluate communicative effectiveness. If assessment practices often prioritize lexical precision, grammatical correctness, and semantic equivalence (while these dimensions remain important), they do not fully capture the complexities of successful communication in multilingual and AI-mediated environments. Students would need to be evaluated on their ability to interpret speaker intentions, convey appropriate levels of politeness, adapt messages to audience expectations, and make contextually appropriate decisions.

Assessment rubrics might therefore include criteria such as pragmatic appropriateness, audience awareness, intercultural sensitivity, communicative effectiveness, and critical evaluation of AI-generated output. In interpreting tasks, assessors could examine how effectively students manage interpersonal relationships, respond to contextual cues, and preserve intended illocutionary force.

Such changes are particularly relevant in Vietnam, where educational traditions have often emphasized examination performance and linguistic correctness. While these priorities have contributed to strong grammatical knowledge, they may not adequately prepare students for the communicative demands of contemporary professional environments. A shift toward communicative and pragmatic assessment would better align T&I training with the realities of globalized, technology-rich workplaces.

Ultimately, the goal of T&I education should not be merely to produce accurate translators and interpreters, but to develop reflective communicators capable of working effectively with both human interlocutors and AI technologies. By integrating pragmatic awareness, authentic learning experiences, critical AI literacy, and communicatively oriented assessment, educators can better prepare students to navigate the complex linguistic and intercultural challenges of the twenty-first century.

Several Assessment Criteria can be suggested as follows:

Dimension	***Descriptor***
Accuracy	Semantic equivalence and linguistic correctness
Pragmatic Force	Recognition and transfer of intent, tone, stance, and implicature
Cultural Appropriateness	Adaptation to sociocultural and communicative norms
Interactional Alignment	Maintenance of interpersonal and communicative relationships
AI Literacy	Ability to evaluate, critique, revise, and justify AI-generated output

These can contribute to T&I education in two important ways. First, it bridges the longstanding gap between theoretical discussions of pragmatics and practical assessment procedures. Second, it can provide a model that can be readily adapted to ESL/EFL classrooms where AI technologies are becoming increasingly accessible. Assessment would focus not only

on accuracy but also on whether students successfully identify and address pragmatic shortcomings. In this way, the students are trained to become reflective communicators capable of integrating linguistic knowledge, cultural and political awareness, and technological competence.

7. Implications for Vietnam and Global Contexts

The arguments presented in this paper have significant implications for both Vietnam and the broader international T&I community. While AI technologies continue to reshape multilingual communication, their successful integration into education and professional practice depends largely on how communicative competence is conceptualized and assessed.

Vietnam represents a particularly compelling case for reform. Over the past two decades, English language education has increasingly embraced communicative approaches; however, in many institutions T&I classroom practices and assessment systems often remain strongly influenced by examination-oriented traditions. The growing availability of AI tools creates both opportunities and challenges within this context.

On the one hand, AI has the potential to democratize access to high-quality language resources. Students from diverse educational backgrounds can now access T&I assistance, conversational practice, interpreting simulations, and immediate feedback at little or no cost. In this sense, AI may function as an educational equalizer, helping bridge gaps in resources and exposure.

On the other hand, uncritical reliance on AI may reinforce surface-level learning and create a false impression of communicative competence. Students who accept machine-generated output without evaluating its pragmatic appropriateness risk overlooking the very skills that distinguish effective translators and interpreters from automated systems. Consequently, AI literacy must be accompanied by pragmatic literacy.

Teacher preparation is equally important. Educators themselves require opportunities to develop confidence in integrating AI tools into teaching and assessment. Professional development programs should therefore address not only technological skills but also strategies for fostering critical thinking, communicative competence, and human-AI collaboration.

Although this discussion draws heavily on the Vietnamese context, the issues raised are globally relevant. Educational institutions worldwide face similar challenges as AI technologies become increasingly embedded in language learning and professional communication. The need to balance technological efficiency with human communicative judgment is therefore not unique to Vietnam but represents a broader concern for translator and interpreter education in the twenty-first century.

8. Conclusion

This paper has argued that the future of T&I education requires a shift from accuracy-centred paradigms toward broader conceptions of communicative competence. While linguistic accuracy remains a necessary component of successful T&I, it is insufficient for ensuring effective communication in increasingly complex and technology-mediated environments.

Drawing on classroom practices, observations, theories of communicative competence, pragmatics, speech acts, translation studies, and AI-assisted communication, we have demonstrated that pragmatic force occupies a central position in T&I training. Meaning is shaped not only by words and grammatical structures but also by intentions, interpersonal relationships, cultural expectations, and contextual factors. Consequently, successful

translators and interpreters must be capable of recognizing and conveying communicative meanings that extend beyond linguistic form.

The analysis has further shown that contemporary AI technologies, despite their remarkable achievements, continue to encounter significant challenges in processing pragmatic dimensions of communication. Neural machine translation systems and large language models can generate fluent and often accurate output, yet they frequently struggle with implicature, politeness, humour, cultural nuance, and context-dependent meaning. These limitations reinforce the continuing importance of human expertise.

AI helps reshape professional roles through human-in-the-loop models of collaboration. In these environments, human practitioners contribute pragmatic judgment, intercultural awareness, ethical responsibility, and communicative intelligence, capacities that remain difficult to replicate technologically. The future of the profession therefore lies not in competition between humans and machines but in productive collaboration between them.

For ESL/EFL contexts such as Vietnam, training programs would need to cultivate pragmatic awareness, critical AI literacy, contextual sensitivity, and strategic decision-making alongside linguistic competence.

Ultimately, communication is a fundamentally human activity. Although technology can facilitate communication, it cannot fully replace the social, cultural, and interpretive processes through which meaning is constructed. As Halliday (1978) suggests through his social semiotic view of language, meaning emerges through human interaction rather than linguistic forms alone. Put simply, meaning is not in words but in people. In an age increasingly shaped by AI, it is precisely this human capacity for understanding, negotiating, and communicating meaning that will remain at the heart of T&I as well as communication.

As a conceptual paper, this study has several limitations, and further research is therefore needed. First, the arguments are primarily derived from own practices, theoretical literature in pragmatics, T&I studies, communicative competence, language assessment, and AI-assisted translation, rather than from original empirical collected data, although the discussion draws on classroom-informed practices and examples from ESL/EFL contexts, particularly Vietnam. These observations were interpreted through a literature-informed thematic analysis rather than a formal empirical data collection process. Consequently, the themes discussed should be viewed as illustrative and exploratory rather than statistically generalizable. Second, the rapidly evolving nature of AI technologies presents an ongoing challenge. Large language models, neural machine translation systems, and speech-processing tools continue to improve at a pace that may outstrip existing pedagogical and assessment frameworks. Consequently, some pragmatic limitations identified here in current AI systems may become less pronounced over time, while new challenges may emerge as we explore the applicability of the framework in diverse linguistic, cultural, and professional contexts.

References

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.463

Bowker, L. (2023). *De-mystifying translation: Introducing translation to non-translators*. Routledge.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/I.1.1

Chapelle, C. A. (2003). *English Language Learning and Technology: Lectures on Applied Linguistics in the Age of Information and Communication Technology*. John Benjamins.

Darong H. C. (2024) 'Teaching speech acts in EFL context', *Education*, 9, doi:10.3389/feduc.2024.1423498.

Dayter D, Locher M. A., and Messerli T. C. (2023) *Pragmatics in Translation: Mediality, Participation and Relational Work*, Cambridge University Press, doi:10.1017/9781009261210.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67–86. https://doi.org/10.1007/BF02504676

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge

House, J. (2015). *Translation Quality Assessment: Past and Present* (2nd ed.). Routledge.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.

Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. In *Proceedings of the First Workshop on Neural Machine Translation* (pp. 28–39). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-3204

Kokorelias, K. M., Metersky. K., Reppas-Rindlisbacher, C., and Al-Hamad, A. (2026). "Beyond

translation: A review into preserving human connection in AI-mediated health language

interpretation", *PLOS Digit Health*, 5(3):1-5, https://doi.org/10.1371/journal.

pdig.0001295

Leech G. N. (1983) *Principles of Pragmatics*, Longman.

Matsushita K. (2026) 'Understanding AI interpreting in context: A comprehension-based evaluation of human vs. machine-generated interpretations in a real-world setting', *International Journal of Language, Translation and Intercultural Communication*, 11:71-85, doi:10.12681/ijltic.44192.

O'Hagan, M. (2020). Introduction: The changing landscape of translation and technology. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 1–16). Routledge.

Pym, A. (2013). Translation skill-sets in a machine-translation age. *Meta, 58*(3), 487–503. https://doi.org/10.7202/1025047ar

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*(2), 91–112. https://doi.org/10.1093/applin/4.2.91

Vieira, L. N. (2019). Post-editing of machine translation. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 319–335). Routledge.

#94 EXPLORING THE INTERPLAY OF MOTIVATION, LEARNER AUTONOMY, AND ACADEMIC PERFORMANCE AMONG ENGLISH-MAJOR STUDENTS AT HUFLIT: A MIXED-METHODS STUDY

Ma Kinh Lac Che[1],[2], Chau Hoang Anh Phan[1], Gia Bao Duong[1] & Minh Quyen Nguyen[1]

Abstract

In Vietnamese tertiary EFL education, students are expected to take responsibility for their own learning while maintaining strong academic outcomes, yet how motivation translates into performance through learner autonomy remains insufficiently examined in institutional settings such as HUFLIT. This mixed-methods study combined survey data from 145 English-major undergraduates with semi-structured interviews from 12 participants. Partial Least Squares Structural Equation Modeling examined proposed relationships, while thematic analysis documented lived experiences. Findings revealed that learner autonomy partially mediated the relationships between both intrinsic and extrinsic motivation and academic performance. Intrinsic motivation had the strongest effect on learner autonomy ($\beta = 0.618$, $t = 12.118$, $p < .001$) and a significant direct effect on academic performance ($\beta = 0.395$, $t = 5.809$, $p < .001$). Learner autonomy was the strongest direct predictor of academic performance ($\beta = 0.436$, $t = 6.141$, $p < .001$). Extrinsic motivation showed significant but weaker effects on learner autonomy ($\beta = 0.244$, $t = 3.342$, $p = .001$) and academic performance ($\beta = 0.123$, $t = 2.085$, $p = .040$). Proportionally, learner autonomy accounted for 40.5% of intrinsic motivation's total effect and 46.3% of extrinsic motivation's total effect. Qualitative analysis identified four themes: (1) internalization of extrinsic motivation, (2) autonomy as a mechanism for addressing academic setbacks, (3) incorporation of informal digital immersion, and (4) resilience strategies for managing motivational setbacks. This study situates Self-Determination Theory and Self-Regulated Learning frameworks within Vietnamese higher education, providing practical insights for cultivating autonomous learning environments that improve English language education outcomes.

Keywords: academic performance, learner autonomy, intrinsic motivation, extrinsic motivation, EFL learners, HUFLIT

1. Introduction

1.1. Background

In Vietnamese tertiary EFL education, students are increasingly expected to take greater responsibility for their own learning while maintaining satisfactory academic outcomes. This expectation is particularly relevant for English-major students, whose language development depends not only on classroom instruction but also on sustained independent practice, strategic resource use, and continuous self-monitoring beyond scheduled lessons. Within this context, motivation and learner autonomy have become central constructs in understanding how students manage their learning and achieve academic progress. Previous research has shown that motivated learners are more likely to persist through difficulty, use effective learning strategies, and engage more deeply with academic tasks (Pintrich & De Groot, 1990;

[1] *Undergraduate student, Faculty of Foreign Languages, HUFLIT*
[2] *Corresponding author; Email: kinhlacma@gmail.com*

Zimmerman & Kitsantas, 2014). From the perspective of Self-Determination Theory, intrinsic motivation reflects students' genuine interest and enjoyment in learning, whereas extrinsic motivation refers to learning driven by external outcomes such as grades, career opportunities, or social expectations (Deci & Ryan, 2000).

In language learning, both forms of motivation may influence performance, but their effects depend on whether students can translate motivation into autonomous learning behaviours, such as setting goals, monitoring progress, seeking resources, and adjusting strategies. Learner autonomy is especially important in EFL contexts because language proficiency develops gradually and requires long-term engagement. Studies in Vietnam suggest that although university students often value independent learning, they may lack the metacognitive strategies needed to practise it effectively (Dang, 2010). Cultural factors such as family expectations, achievement pressure, and career-oriented learning may also shape how Vietnamese students internalize motivation and develop autonomy (Truong, 2025). The relationship among motivation, learner autonomy, and academic performance requires further examination in specific institutional contexts such as HUFLIT, where English-major students are expected to develop both academic competence and independent learning capacity.

1.2. Problem Statement

1.2.1. Practical Problem

Third-year English-major students at HUFLIT confront graduation requirements that impose substantial independent learning demands. As an institution that combines foreign language training with information technology education, it imposes graduation requirements that are particularly demanding: English-major students must achieve Microsoft Office Specialist (MOS) certification and a minimum of IELTS 7.0 or TOEIC 780. Beyond these certification thresholds, the university explicitly expects considerable self-directed study outside scheduled contact hours. In this setting, independent learning functions as a structural precondition for academic progression. The institution, however, possesses no evidence base delineating which motivational conditions most effectively cultivate the autonomous behaviour that its requirements presuppose. Instructional decisions concerning how best to support self-directed learning are made on institutional intuition, not empirical evidence. The practical problem is twofold: HUFLIT enforces rigorous self-directed learning demands on its English majors yet lacks a corresponding framework for understanding the motivational antecedents of the autonomous conduct those demands necessitate.

1.2.2. Research Gap

The literature to date, while extensive, has not addressed these dynamics in the institutional and cultural context that HUFLIT represents.

Western universities have supplied the samples for most published work on motivation and autonomy in language learning (Noels et al., 2000; Garcia & Pintrich, 1996). Assessment practices, pedagogical norms, and the cultural meanings attached to academic success in those contexts differ in ways that limit straightforward generalization to Vietnamese higher education. Whether findings obtained with Western undergraduates can illuminate the experiences of English majors at a Vietnamese institution remains an open empirical question.

Prior studies have treated motivation and autonomy as separate lines of inquiry which lead to a further gap in methodology approach. No investigation has positioned both of these factors simultaneously within a single structural model that estimates direct effects, indirect

pathways, and mediated relationships in one analytic framework. The field has also relied predominantly on single-method survey designs, which capture statistical associations but shed little light on the subjective processes through which students navigate motivational change or develop autonomous habits (Chong & Reinders, 2025).

Vietnamese scholarship provides a foundation but not a precedent. Research within the country has confirmed the importance of self-learning in general higher education (Dang, 2010; Dao et al., 2025). None of this work, however, has examined third-year English majors at an institution where self-directed learning is structural requirement rather than a pedagogical preference. The present study was designed to address these three gaps.

1.3. Aim and Objectives

This study aims to investigate the relationships among intrinsic motivation, extrinsic motivation, learner autonomy, and academic performance among third-year English-major students at HUFLIT. It also examines whether learner autonomy mediates the relationship between motivation and academic performance and explores how students' motivational and autonomous learning experiences shape their academic development. The specific objectives are:

To identify the reported levels of intrinsic motivation, extrinsic motivation, learner autonomy, and academic performance among third-year English-major students at HUFLIT.

To examine the effects of intrinsic and extrinsic motivation on learner autonomy.

To examine the direct effects of intrinsic motivation, extrinsic motivation, and learner autonomy on academic performance.

To test the mediating role of learner autonomy in the relationships between intrinsic motivation, extrinsic motivation, and academic performance.

To explore students' experiences of motivational change, autonomous learning practices, and responses to academic difficulties.

1.4. Research Questions

Three research questions organized the investigation:

RQ1: What are the reported levels of intrinsic motivation, extrinsic motivation, learner autonomy, and academic performance among HUFLIT English-major students?

RQ2: How do intrinsic motivation, extrinsic motivation, and learner autonomy relate to academic performance among HUFLIT English-major students?

RQ3: To what extent does learner autonomy mediate the relationship between motivation and academic performance?

1.5. Significance

Theoretically, the study contributes to the literature in two respects: it extends Self-Determination Theory and Self-Regulated Learning frameworks to a Vietnamese EFL context where they have been widely cited but rarely tested with simultaneous structural modeling and qualitative corroboration, and it responds to a documented methodological gap.

The practical significance of the study is grounded in a specific institutional need. HUFLIT requires English majors to complete substantial self-directed learning beyond timetabled hours and to meet external certification benchmarks, including IELTS 7.0 or TOEIC 780, for graduation. Despite these structural expectations, the university has lacked an evidence base identifying which motivational conditions most effectively produce autonomous behaviour and improved academic outcomes.The findings supply this evidence by establishing that intrinsic motivation drives learner autonomy at more than twice the strength of extrinsic motivation and that autonomy is itself the strongest direct predictor of performance. These results carry actionable implications for HUFLIT: investment in autonomy-supportive teaching, curriculum structures that scaffold self-directed learning, and institutional acknowledgment of informal digital immersion are empirically warranted, not aspirational.

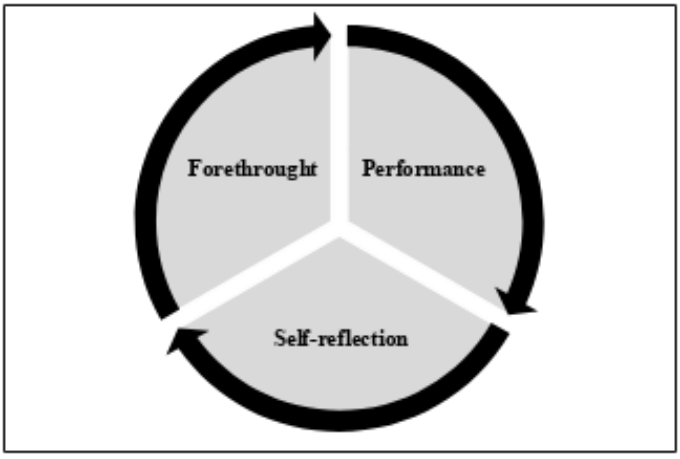
2. Literature Review

2.1. Theoretical Framework

How academic outcomes come to be shaped by motivation is a question that cannot be adequately addressed through a simple listing of variables; what is required instead is an account of the psychological mechanisms by which a learner's drive is converted into the kind of sustained effort that produces measurable achievement. This investigation has been structured around two principal theoretical traditions. Self-Determination Theory (SDT) has been drawn upon to explain where motivation originates and under what conditions it becomes internalized, while Self-Regulated Learning (SRL) theory has been employed to describe the behavioural processes through which motivation is translated into academic attainment. Where the specific circumstances of Vietnamese EFL education raise questions that are not fully addressed by these two frameworks alone, Expectancy-Value Theory and Goal Orientation Theory have been used to supplement the analysis.

Figure 1

Cyclic Phase Model of SRL (Zimmerman, 2002)



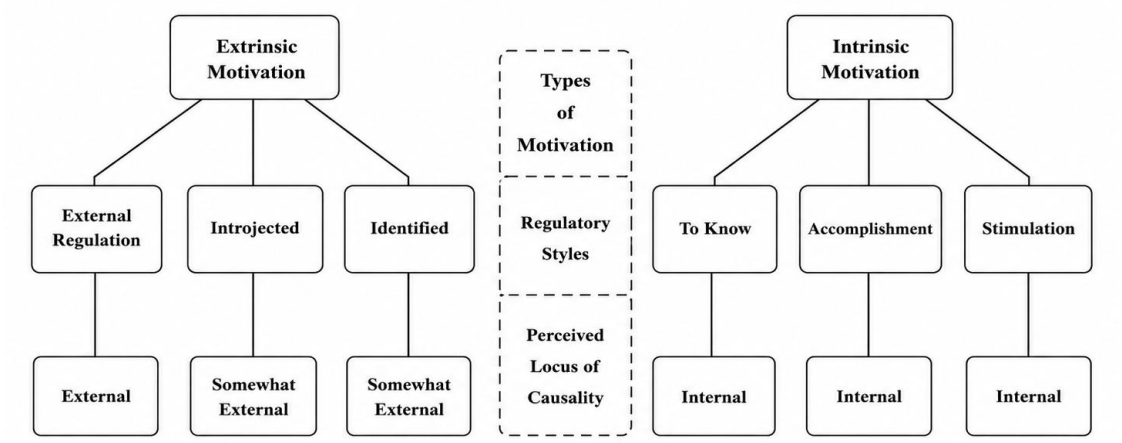
The self-regulatory process was characterized by Zimmerman (2002) as a cycle composed of three distinct phases, each of which feeds into the next and is characterized by its own set of cognitive activities. During the forethought phase, goals are established by learners and strategies are selected on the basis of what is judged to be most appropriate for the task at hand. In the performance phase, the execution of those strategies is monitored against the goals that were set, and cognitive and metacognitive techniques are deployed as circumstances

require. The self-reflection phase is where outcomes are evaluated and inefficiencies are identified, after which the cycle is repeated. What is emphasized by this model, in contrast to more passive conceptions of learning, is the position of the learner as an active agent who continuously calibrates their effort in response to the feedback generated by their own monitoring activities. It has been confirmed by Sinkkonen and Tapani (2024) that this cyclical framework continues to occupy a central position in SRL research conducted within higher education settings. It was also noted by these authors that digital learning environments have introduced challenges for which the original model was not designed, including the need to manage online resources, navigate content that is shaped by algorithmic curation, and maintain concentration in contexts that have been engineered to compete for attention.

The relevance of self-regulation to language learning has been difficult to overstate in the literature, given the manner in which proficiency tends to accumulate. Because language development proceeds slowly and unevenly, and because the demands of different tasks can vary considerably, sustained engagement over extended periods is required of learners, along with a capacity to adjust their strategies as those demands shift (Wenden, 1998). Students who make habitual use of metacognitive strategies, maintain a robust sense of efficacy, and engage in the systematic tracking of their own progress have been found to outperform those who depend on surface-level study routines (Pintrich & De Groot, 1990; Zimmerman & Kitsantas, 2014). Within formal education, self-regulation has been conceptualized as something more than a study skill; it has been positioned as a foundational mechanism through which autonomous capacity is developed, by which learners are furnished with the metacognitive infrastructure needed to direct their own progress without ongoing teacher supervision (Nakata, 2014).

Figure 2

Motivation types, regulatory styles, and perceived locus of causality



A distinction that has become central to the study of academic motivation was drawn by Deci and Ryan (2000) between behaviour that is undertaken for the inherent satisfaction it provides, which was classified as intrinsically motivated, and behaviour that is driven by outcomes separable from the activity itself, such as grades, credentials, or social recognition, which was classified as extrinsically motivated. The value of SDT for the present investigation, however, lies less in this binary distinction than in how externally regulated behaviour has been mapped by the theory along a continuum of internalization. At one end of this continuum, external regulation refers to behaviour that is maintained through the administration of rewards

or the avoidance of punishment. Further along, introjected regulation has been characterized as involving internal pressure that is rooted in guilt or self-esteem. Identified regulation is said to emerge when a person has come to consciously value an activity as personally important, and integrated regulation — positioned as the most autonomous form of extrinsic motivation — is understood to occur when externally originating values have been fully assimilated into the self, so that the behaviour in question is experienced as self-determined rather than externally imposed. The same activity, whether it involves preparing for an examination or completing a course assignment, can be powered by any of these regulatory processes, and the psychological outcomes associated with each have been shown to differ considerably.

The conditions under which internalization occurs have been specified by SDT with reference to three basic psychological needs such as autonomy, competence, and relatedness, which are understood to function as catalysts when they are satisfied (Deci & Ryan, 2000). The available evidence has been broadly consistent with this proposition. Among L2 learners, intrinsic motivation and identified regulation have been found to be predictive of effort, persistence, and actual achievement (Noels et al., 2000). Sustained gains in both engagement and academic performance have been observed when autonomy-supportive teaching practices are implemented (Reeve & Jang, 2006). Similar findings have been reported across a range of foreign language contexts (Dornyei & Csizer, 2012), and positive associations between perceived classroom autonomy and intrinsic goal orientation, task value, and self-efficacy have been identified among college students (Garcia & Pintrich, 1996).

Two additional theoretical perspectives have been used to connect these frameworks to the Vietnamese higher education context within which the present study was conducted. According to Expectancy-Value Theory (EVT), academic behaviour is shaped by students' expectations regarding their likelihood of success and the value that is attached by them to the task in question (Eccles & Wigfield, 2020). This has been considered particularly relevant in settings where English proficiency has come to function as a gateway to employment, family approval, and academic advancement (Truong, 2025). Goal Orientation Theory (GOT) has been used to distinguish between a mastery orientation, in which the development of competence is the primary objective, and a performance orientation, in which the demonstration of ability relative to others is prioritized (Pintrich & Schunk, 2002). Both perspectives have contributed to the interpretation of the data collected for this study, though neither has been assigned the central explanatory role that SDT and SRL occupy.

2.2. Prior Studies and Hypothesis Development

The theoretical traditions reviewed in Section 2.1 converge on a central proposition: motivated learners who take ownership of their learning tend to achieve more. Yet translating this general insight into a set of testable, directional predictions requires a closer look at what the empirical record actually says about each link in the chain. The evidence is not uniformly strong across all pathways, and some relationships rest on firmer ground than others. The pages that follow examine the seven relationships that together form the conceptual model tested in this study, moving from the most established connections to those for which the literature offers more qualified support.

2.2.1. Intrinsic Motivation and Academic Performance

Documentation of the positive effects of intrinsic motivation on educational outcomes existed across a wide range of settings and age groups. Students who engage with learning

material (in the context of English material) out of genuine interest, curiosity, and enjoyment tend to grasp the concept at greater depth, show persistence when faced difficulty, and deploy more sophisticated cognitive and metacognitive strategies than those whose efforts are primarily for external goals (Deci & Ryan, 2000). Early research in the classroom environment by Pintrich and De Groot (1990) showed that intrinsic value was among the strongest motivational predictors of achievement, operating partly through its influence on self-regulated strategy use. A similar pattern has been reported in the domain of second language acquisition, where intrinsically motivated learners consistently demonstrate superior effort, persistence, and attainment. Noels et al. (2000) discovered that intrinsic motivation and more self-determined forms of regulation predicted both subjective engagement and objective language outcomes. More recently, Zhou and Ahmad (2025) confirmed that learning motivation exerted a positive and significant impact on academic outcomes. At the meta-analytic level, evidence across multiple studies was synthesised by Alamer et al. (2025) and concluded that autonomous forms of motivation (intrinsic motivation and identified regulation) correlated consistently with L2 achievement, whereas controlled motivation did not. When these findings are considered alongside Zimmerman and Kitsantas's (2014) evidence that self-motivated students outperform peers who rely on external incentives, a clear empirical trend emerges: the more a student's engagement is sourced from pure interest in the subject, the stronger the academic results tend to be. Based on this evidence, the first hypothesis is proposed:

H1: Intrinsic motivation has a positive direct effect on academic performance.

2.2.2. Extrinsic Motivation and Academic Performance

Motivation that originates outside the learner's interest demonstrates a more complicated mechanism. Extrinsically regulated behaviour is directed by outcomes that are separable from the activity itself such as grades, career credentials, social approval, or the avoidance of negative consequences. SDT does not dismiss extrinsic motivation as inherently detrimental; rather, it maps extrinsic regulation onto a continuum that ranges from fully controlled compliance to fully integrated self-endorsement (Deci & Ryan, 2000). Where a student falls on this continuum matters considerably for outcomes. At the controlled end, learning tends to be shallow and engagement fragile. Vallerand and Bissonnette (1992) demonstrated that less self-determined motivational styles predicted weaker persistence and lower academic performance in a prospective study of college students.

Internalization, however, alters the dynamic considerably. When a student encounters a graduation requirement (a minimum of IELTS 7.0 or TOEIC 780 and MOS certificates) or perceives a career credential as personally meaningful, extrinsic motivation can support achievement in ways that resemble intrinsic motivation, albeit with smaller effect sizes. Nagle (2021), applying EVT to university-level Spanish learners, found that utility value, an inherently extrinsic orientation, significantly predicted both course persistence and final grades. Zhou and Ahmad (2025) reported a similar pattern: extrinsic motivation made a positive contribution to academic performance, despite having a substantially smaller coefficient than intrinsic motivation. The weight of evidence therefore suggests a direct pathway from extrinsic motivation to performance, but one that is narrower than the intrinsic route:

H2: Extrinsic motivation has a positive direct effect on academic performance.

2.2.3. Intrinsic Motivation and Learner Autonomy

Among the four constructs in the present model, the bond between intrinsic motivation and learner autonomy has attracted perhaps the most sustained theoretical attention. SDT

conceptualizes autonomy as a fundamental psychological need. Activities pursued for their inherent satisfaction should, by the theory's own logic, orient learners toward greater self-direction (Deci & Ryan, 2000). In classroom settings, this bidirectional dynamic operates when students perceive genuine choice in their academic work, they report stronger intrinsic motivation, and students who are intrinsically motivated actively seek out opportunities for independent learning. Garcia and Pintrich (1996) provided early evidence for this reciprocal pattern among college students, showing positive associations between perceived classroom autonomy and intrinsic goal orientation, task value, and self-efficacy.

More recent work has sharpened the picture considerably. Ceylan (2021) reported a significant positive correlation between learner autonomy and motivation in a sample of Turkish EFL students, with the relationship holding across multiple dimensions of autonomous learning. The argument was taken a step further through experimental manipulation: when students were given the freedom to choose whether their attendance counted toward their grade, they attended more reliably and performed better than students for whom attendance had been mandated, showing that genuine choice operates as the mechanism driving autonomous engagement (Cullen & Oppenheimer, 2024). Wang et al. (2024) extended this reasoning by demonstrating that teacher autonomy support enhanced Chinese EFL undergraduates' speaking performance indirectly, operating through satisfaction of basic psychological needs and increased classroom engagement.

Evidence from prior studies in Vietnam points in the same direction. Dang (2010) observed that English-major students who valued the language for its intrinsic properties were far more likely to engage in self-directed study than those driven solely by examination pressures or family expectations. Thi et al. (2022) surveyed EFL students across Vietnam and reached a similar conclusion that the strongest predictor of autonomous learning was whether students studied by choice, not whether they were chasing a grade. In addition, collectivist values cut both ways when students have internalized them, autonomy grows and when students comply without personal endorsement, autonomy suffers (Linh & Loi, 2024). Taken together, these findings support a strong, positive pathway from intrinsic motivation to learner autonomy. Hence:

H3: Intrinsic motivation has a positive direct effect on learner autonomy.

2.2.4. Extrinsic Motivation and Learner Autonomy

At first glance, the notion that externally driven motivation might enhance autonomous learning seems less plausible than the intrinsic case. Nevertheless, it follows logically from the internalization continuum described in SDT. When students operate under external pressures such as demands of a competitive job market, certification threshold, or even strong family expectations, they are not simply passive. These conditions can elicit behaviours that autonomy scales measure: systematic goal-setting, strategic resource use, and persistent self-monitoring. Deci and Ryan (2000) argued that satisfaction of the needs for competence and relatedness, alongside autonomy, facilitates the gradual transformation of externally regulated behaviour into personally endorsed action. Empirical support for this proposition, though less voluminous than for the intrinsic pathway, has accumulated steadily. Tao and Jing (2023) found that autonomous motivation served as a significant mediator between self-discipline and reduced academic procrastination, suggesting that behaviours initially sustained by external discipline can, over time, become internalized and self-regulated. Danilenko et al. (2018) contended that the structural features of contemporary higher education, such as modular curricula, independent project work, and digital learning platforms, create conditions under which externally prompted study gradually evolves into autonomous practice. The Vietnamese data

offer a parallel trajectory. Several studies have documented that although undergraduates initially study for instrumental reasons, many develop genuine autonomous learning routines as they progress through their degree programmes (Dang, 2010; Dao et al., 2025). None of this implies that the extrinsic pathway rivals the intrinsic one in magnitude; motivation that remains contingent on external rewards without adequate internalisation tends to yield compliance rather than authentic self-regulation. The fourth hypothesis therefore captures the more qualified prediction:

H4: Extrinsic motivation has a positive direct effect on learner autonomy.

2.2.5. Learner Autonomy and Academic Performance

The SRL framework describes the process of converting energy from motivation into measurable achievement through the mechanism provided by autonomy. In detail, autonomous learners move through phases of forethought, performance, and self-reflection (Zimmerman & Kitsantas, 2014). Language learning follows a protracted and uneven developmental trajectory; it places particularly heavy demands on this cycle, making self-regulatory competence a near-necessity for sustained improvement. The empirical case for a direct autonomy-to-performance link has been built across multiple EFL contexts. Cao and Pho (2025) documented a significant positive relationship between learner autonomy and academic performance in a tertiary EFL sample, with autonomous learners registering notably higher grades than their less self-directed peers. Nakata (2014) argued that self-regulation constitutes a foundational mechanism for developing autonomous capacity in formal schooling, precisely because it equips learners with the metacognitive tools to manage their own progress independent of teacher direction. Intervention-based studies have strengthened this case further. Ismail et al. (2023) showed that embedding self-regulated learning practices into EFL assessment design boosted both learner autonomy and self-efficacy, with downstream gains in achievement. Hsieh and Hsieh (2019) observed a similar pattern at the behavioural level: students who independently curated resources and managed out-of-class study outperformed those who relied exclusively on instructor-provided materials. In a setting such as HUFLIT, where students are expected to complete substantial self-directed work outside timetabled hours and where graduation hinges on externally benchmarked proficiency thresholds, autonomy is unlikely to be merely helpful; it is plausibly the single most important behavioural determinant of academic outcomes. The following hypothesis captures this expectation:

H5: Learner autonomy has a positive direct effect on academic performance.

2.2.6. Learner Autonomy as a Mediator of Intrinsic Motivation

The five direct relationships reviewed above, taken together, point toward a more complex structure than a simple additive model would suggest. Intrinsic motivation does not appear to influence academic performance through a single, unbroken channel. Part of its effect travels through the autonomous behaviours that intrinsic interest tends to generate. A student who genuinely enjoys English does not merely pay closer attention in lectures; they also set personal learning goals, seek out supplementary materials, practice outside the classroom, and evaluate their own development, each of which independently strengthens academic outcomes. The logic is straightforward: intrinsic interest fuels self-directed action, and self-directed action produces achievement.

Support for this mediated pathway comes from the pattern of evidence already assembled. Section 2.2.3 established that intrinsic motivation robustly predicts learner autonomy; Section 2.2.5 established that learner autonomy robustly predicts academic performance; and Section 2.2.1 confirmed that a direct route from intrinsic motivation to

performance also exists. The mediation hypothesis simply formalises what these three findings jointly entail, namely, that autonomy carries a meaningful share of intrinsic motivation's total effect. Accordingly:

H6: Learner autonomy mediates the relationship between intrinsic motivation and academic performance.

2.2.7. Learner Autonomy as a Mediator of Extrinsic Motivation

A parallel mediating logic extends to extrinsic motivation, though the mechanism differs in a theoretically important manner. When a student's primary drive is instrumental, whether a required certification, a desired career path, or a family expectation, that drive alone cannot elevate academic performance unless it translates into concrete, self-regulated activity. A student who needs IELTS 7.0 to graduate does not achieve that score simply because the requirement exists; they achieve it because the requirement prompts them to build a study schedule, seek feedback on their writing, track their listening comprehension over time, and adjust their strategies when practice test results disappoint. If they fail to develop these autonomous behaviours, the external pressure remains inert. The empirical basis for this expectation follows the same logic that underpins **H6**.

Extrinsic motivation positively predicts learner autonomy (Section 2.2.4), learner autonomy positively predicts academic performance (Section 2.2.5), and the direct, unmediated effect of extrinsic motivation on performance is comparatively weak (Section 2.2.2). Because externally sourced motivation relies more heavily on behavioural implementation than the intrinsic variety does, and no inherent enjoyment carries the learner through difficulty, a larger share of its total effect should pass through the autonomy pathway. The final hypothesis is therefore:

H7: Learner autonomy mediates the relationship between extrinsic motivation and academic performance.

2.3. Hypotheses

Based on the theory and evidence reviewed in Sections 2.1 and 2.2, seven hypotheses were formulated. Figure 3 presents the conceptual framework with these hypothesized relationships:

H1: Intrinsic motivation has a positive direct effect on academic performance.

H2: Extrinsic motivation has a positive direct effect on academic performance.

H3: Intrinsic motivation has a positive direct effect on learner autonomy.

H4: Extrinsic motivation has a positive direct effect on learner autonomy.

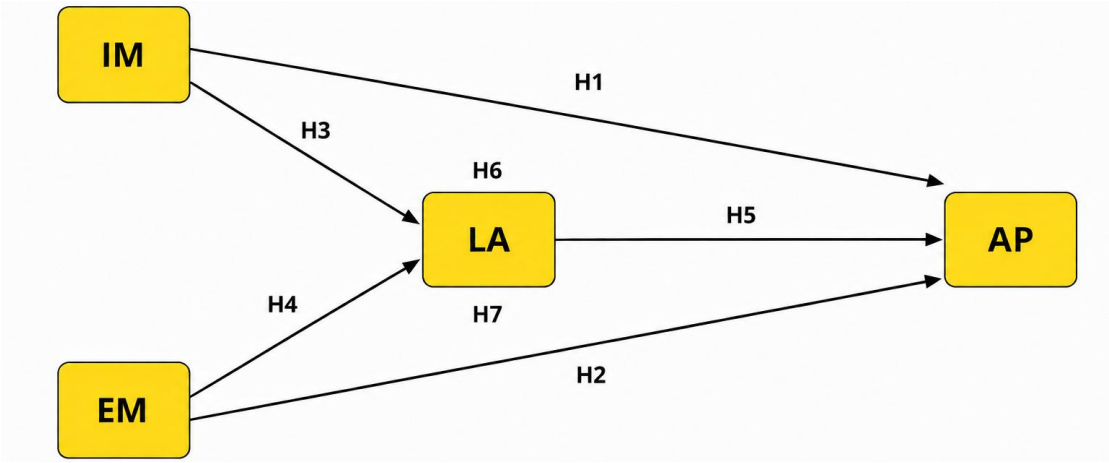
H5: Learner autonomy has a positive direct effect on academic performance.

H6: Learner autonomy mediates the relationship between intrinsic motivation and academic performance.

H7: Learner autonomy mediates the relationship between extrinsic motivation and academic performance.

Figure 3

Conceptual Framework of Motivation, Learner Autonomy, and Academic Performance



3. Methodology

3.1. Research Design

The study employed a mixed-methods explanatory sequential design (Creswell & Plano Clark, 2018). Quantitative data were collected and analyzed first, followed by qualitative data to elaborate and contextualise the quantitative results. The rationale is twofold: quantitative methods identify patterns, test structural relationships, and estimate effect magnitudes, while qualitative methods reveal the subjective experiences and contextual nuances underlying statistical associations.

3.2. Research Setting and Participants

The study was conducted at HUFLIT during the second semester of the 2025 academic year. A convenience sampling method was used to recruit participants from the population of third-year students enrolled in English-major programmes within the Faculty of Foreign Languages. Third-year students were targeted because they had completed sufficient exposure to the university's English-language curriculum to have formed relatively stable motivational orientations and autonomous learning patterns, while still remaining actively engaged in major coursework. Within this target group, participants were recruited on a voluntary basis from accessible class sections.

As shown in Table 1, most participants were female (n = 95, 65.5%). English Language Teaching represented the largest major group (n = 60, 41.4%), followed by Business English (n = 40, 27.6%). Most students reported a GPA of 2.5 or above (n = 140, 96.6%), indicating that higher-performing students may have been more likely to participate in the survey.

Table 10

Demographic Characteristics of Participants (N = 145)

Demographic	Variable	Frequency	Percentage (%)
Gender	Male	50	34.5
	Female	95	65.5
Major	English Language Teaching	60	41.4
	Business English	40	27.6

Demographic	Variable	Frequency	Percentage (%)
	Translation & Interpretation	15	10.3
	English-Chinese Bilingual	15	10.3
	Office Administration	10	6.9
	Logistics	5	3.4
GPA (Self-reported)	< 2.5	5	3.4
	2.5–3.0	58	40.0
	3.0–3.2	42	29.0
	3.2–4.0	40	27.6

The qualitative sample consisted of twelve participants selected through purposive maximum-variation sampling to ensure representation across the range of motivational profiles (high intrinsic, high extrinsic, mixed, low-moderate, and transition from extrinsic to intrinsic) and GPA levels (2.5–3.0, 3.0–3.5, and 3.5–4.0). Pseudonyms were assigned to protect participant confidentiality.

Table 11

Interview Participant Characteristics

Pseudonym	Major	GPA Range	Motivation Profile
January	Business English	3.5–4.0	High Intrinsic
February	English Language Teaching	3.5–4.0	High Intrinsic
March	Translation & Interpretation	3.0–3.5	Mixed (Internalizing)
April	English-Chinese Bilingual	3.0–3.5	High Extrinsic
May	Business English	2.5–3.0	Low-Moderate
June	English Language Teaching	3.5–4.0	High Intrinsic
July	Office Administration	3.0–3.5	Moderate Extrinsic
August	Business English	3.5–4.0	High Intrinsic
September	Translation & Interpretation	3.0–3.5	Mixed
October	English Language Teaching	3.5–4.0	Transition (Extrinsic→Intrinsic)
November	Business English	3.0–3.5	High Intrinsic
December	English-Chinese Bilingual	3.5–4.0	High Intrinsic

Note. Pseudonyms are used to protect participant confidentiality.

3.3. Research Instruments

A structured questionnaire comprising 24 items was developed to measure Intrinsic Motivation, Extrinsic Motivation, Learner Autonomy, and Academic Performance, with six items for each construct. The motivation items were adapted from Self-Determination Theory and L2 motivation research (Deci & Ryan, 2000; Noels et al., 2000; Vallerand & Bissonnette, 1992; Afzal et al., 2010). Learner autonomy items were informed by learner autonomy and

self-regulated learning studies (Dang, 2010; Gardner, 2007; Nakata, 2014; Zimmerman & Kitsantas, 2014), while academic performance items were adapted from studies on motivation, self-regulation, and achievement (Pintrich & De Groot, 1990; Alhadabi & Karpinski, 2020; Cao & Pho, 2025). All items were contextualized for Vietnamese tertiary EFL students and rated on a 5-point Likert scale. The questionnaire was piloted with 10 HUFLIT English-major students excluded from the primary sample. Cronbach's alpha values from the pilot confirmed satisfactory internal consistency for all four scales. Experts in Applied Linguistics and TESOL reviewed the items for content validity, and their feedback informed refinements to item wording for the target population.

Semi-structured interviews lasting 8 to 15 minutes were conducted with the twelve qualitative participants, either in person or via digital conferencing platforms, and were audio-recorded with participants' consent. The semi-structured interview protocol was developed based on the study's conceptual framework and prior research (Deci & Ryan, 2000; Noels et al., 2000; Dang, 2010; Gardner, 2007; Zimmerman, 2002). It focused on four areas: students' initial motivation for choosing English, factors sustaining or weakening motivation, autonomous learning behaviours, and strategies for overcoming motivational difficulties. Credibility was supported through member-checking, in which three participants reviewed summaries of their interview responses and confirmed the accuracy of the researchers' interpretations.

3.4. Data Collection Procedures

Data collection proceeded in three phases: distributing the Google Forms survey to eligible third-year English majors via Zalo groups and QR codes (open for two weeks with one reminder); inviting survey respondents to volunteer for interviews; and selecting twelve participants via maximum-variation sampling for individual interviews. All responses were anonymized.

3.5. Data Analysis

Quantitative data were analyzed in two stages. Descriptive statistics, including means, standard deviations, and frequency distributions, were computed using SPSS to characterize the sample's motivational and autonomy profiles. Partial Least Squares Structural Equation Modeling was then conducted using SmartPLS 4.0. PLS-SEM was selected over covariance-based SEM because of its suitability for predictive and exploratory research, its robustness with smaller samples, and its capacity to accommodate non-normal data distributions (Hair et al., 2019).

The PLS-SEM analysis followed the two-stage procedure recommended by Hair et al. (2019). In the first stage, the measurement model was evaluated through indicator reliability (outer loadings $\geq .7$), internal consistency reliability (Cronbach's alpha and composite reliability $\geq .7$), convergent validity (AVE $\geq .5$), and discriminant validity (HTMT $< .9$). In the second stage, the structural model was assessed through collinearity diagnostics (VIF < 3), path coefficients and significance testing (bootstrapping with 5,000 subsamples), explanatory power (R^2), effect sizes (f^2), and mediation analysis using the bootstrapped indirect effects procedure.

Qualitative data were analyzed following Braun and Clarke's (2006) thematic analysis procedures. Interview recordings were transcribed and read repeatedly to ensure familiarization

with the data. Initial codes were generated inductively from participants' responses, then grouped into candidate themes that reflected recurring patterns in students' motivational changes, autonomous learning behaviours, and coping strategies. These themes were reviewed, refined, and confirmed through member-checking to enhance credibility.

3.6. Reliability and Validity

Reliability and validity were addressed through multiple procedures. The questionnaire items were adapted from established studies on motivation, learner autonomy, self-regulated learning, and academic performance, then contextualized for Vietnamese tertiary EFL students. Content validity was strengthened through expert review in Applied Linguistics and TESOL, and wording was refined after a pilot test with 10 students. Internal consistency was assessed using Cronbach's alpha, composite reliability, and rho_A, while convergent validity was examined through outer loadings and Average Variance Extracted (AVE). Discriminant validity was evaluated using the HTMT criterion. For the qualitative phase, credibility was enhanced through repeated reading, thematic refinement, and member-checking.

4. Results

4.1. Quantitative Findings

4.1.1. Descriptive Statistics

All four variables demonstrated high mean scores, indicating that HUFLIT English-major students generally reported positive levels of motivation, learner autonomy, and perceived academic performance. Table 3 presents the descriptive statistics for all constructs.

Table 12

Descriptive Statistics for All Variables (N = 145)

Variables	***M***	***SD***
Intrinsic Motivation	3.95	0.83
Extrinsic Motivation	3.85	0.79
Learner Autonomy	3.97	0.76
Academic Performance	3.95	0.77

Note. M = Mean; *SD* = Standard Deviation

As shown in Table 3, Learner Autonomy had the highest mean score (M = 3.97, SD = 0.76), followed closely by Intrinsic Motivation and Academic Performance, both at M = 3.95. Extrinsic Motivation was slightly lower but still above the midpoint of the 5-point scale (M = 3.85, SD = 0.79). At the item level, several responses help clarify these construct-level patterns. For intrinsic motivation, the strongest item was IM4: "I genuinely love learning the English language" (M = 4.16, SD = 0.856), indicating strong personal interest in English. For extrinsic motivation, EM2: "I learn English because it will help me get a better job in the future" received the highest rating (M = 4.14, SD = 0.933), showing that career value was the most salient external driver. Within learner autonomy, LA5: "I take responsibility for my English learning because it matters to me" had the highest score (M = 4.09, SD = 0.83), suggesting that students generally perceived themselves as responsible for their own learning. For academic performance, AP1: "My English course grades reflect my learning progress" was the highest-rated item (M = 3.99, SD = 0.901). Overall, the descriptive findings indicate that students showed strong intrinsic interest, clear career-oriented motivation, a high sense of learning responsibility, and generally positive perceptions of their academic progress.

4.1.2. Measurement Model Assessment

Prior to testing the structural hypotheses, the measurement model was evaluated against four criteria established by Hair et al. (2019): indicator reliability, internal consistency reliability, convergent validity, and discriminant validity.

Overall, the measurement model demonstrated satisfactory reliability and validity. All indicators loaded strongly on their intended constructs, internal consistency was confirmed through Cronbach's alpha, rho_A, and composite reliability, and convergent validity was supported by AVE values above the recommended threshold. Discriminant validity was also established, as all HTMT values remained below 0.90. These results indicate that the four constructs: Intrinsic Motivation, Extrinsic Motivation, Learner Autonomy, and Academic Performance were measured reliably and distinctly, providing a sound basis for testing the structural model.

Indicator reliability. All 24 indicators produced outer loadings exceeding the 0.7 threshold. Loadings ranged from 0.731 for EM6 to 0.907 for IM2. Table 4 displays the complete loading matrix.

Table 13

Outer Loadings for All Indicators

	AP	**EM**	**IM**	**LA**
AP1	0.767			
AP2	0.802			
AP3	0.877			
AP4	0.894			
AP5	0.872			
AP6	0.848			
EM1		0.733		
EM2		0.775		
EM3		0.823		
EM4		0.752		
EM5		0.805		
EM6		0.731		
IM1			0.890	
IM2			0.907	
IM3			0.834	
IM4			0.873	
IM5			0.878	
IM6			0.824	
LA1				0.853
LA2				0.871
LA3				0.849
LA4				0.812
LA5				0.803
LA6				0.810

Note. All outer loadings exceed the recommended threshold of 0.7 Hair et al. (2019).

The consistently high loadings confirm that each observed variable adequately represents its respective latent construct.

Internal consistency and convergent validity. Table 5 presents Cronbach's alpha (α), composite reliability (rho_a and rho_c), and Average Variance Extracted (AVE) for all four constructs.

Table 14

Measurement Model Summary

Construct	α	rho_a	rho_c	AVE
Intrinsic Motivation	0.933	0.935	0.948	0.754
Extrinsic Motivation	0.865	0.871	0.897	0.594
Learner Autonomy	0.912	0.916	0.932	0.695
Academic Performance	0.919	0.922	0.937	0.713

Note. α = Cronbach's alpha; rho_c = composite reliability; AVE = average variance extracted.

Cronbach's alpha values ranged from 0.865 (Extrinsic Motivation) to 0.933 (Intrinsic Motivation), all substantially exceeding the 0.7 threshold. Composite reliability (rho_c) values ranged from 0.897 to 0.948. Three of the four constructs cleared the 0.9 threshold commonly associated with excellent reliability; Extrinsic Motivation landed in the good range, which is not surprising given its broader conceptual umbrella. All AVE values exceeded the 0.5 benchmark established by Fornell and Larcker (1981). Intrinsic Motivation had the highest AVE (0.754), indicating that this latent construct explains 75.4% of the variance in its six indicators. The comparatively lower AVE for Extrinsic Motivation (0.594) is theoretically consistent with the broader conceptual scope of this construct, which spans diverse external influences including grades, career prospects, parental expectations, graduation requirements, social recognition, and family obligation.

Discriminant validity. The Heterotrait-Monotrait (HTMT) ratio of correlations, presented in **Table 6**, confirmed discriminant validity for all construct pairs, with all values falling below the 0.9 threshold recommended by Henseler et al. (2015).

Table 15

HTMT Ratio Matrix for Discriminant Validity

Construct	AP	EM	IM	LA
Academic Performance (AP)	–			
Extrinsic Motivation (EM)	0.612	–		
Intrinsic Motivation (IM)	0.847	0.598	–	
Learner Autonomy (LA)	0.875	0.601	0.823	–

Note. All HTMT values < 0.9 Henseler et al. (2015).

The highest HTMT was between Academic Performance and Learner Autonomy (0.875), expected given their structural relationship, and the lowest between Extrinsic and Intrinsic Motivation (0.598), confirming their empirical distinction.

4.1.3. Structural Model Assessment

Collinearity among predictor constructs was evaluated; all inner VIF values fell well below the threshold of 3, confirming that multicollinearity does not bias the regression estimates (Table 7).

Table 16

Inner VIF Values for Predictor Constructs

Endogenous Construct	Predictor	VIF	Status
Learner Autonomy	Intrinsic Motivation	1.456	No collinearity
Learner Autonomy	Extrinsic Motivation	1.456	No collinearity
Academic Performance	Intrinsic Motivation	1.523	No collinearity
Academic Performance	Extrinsic Motivation	1.489	No collinearity
Academic Performance	Learner Autonomy	1.612	No collinearity

Note. VIF < 3 indicates no problematic multicollinearity Hair et al. (2019).

The structural model was tested using bootstrapping with 5,000 subsamples. Table 8 presents the standardised path coefficients, standard errors, *t*-values, *p*-values, and 95% confidence intervals for all five direct-effect hypotheses.

Table 17

Structural Model Path Coefficients and Hypothesis Testing

H	Path	β	*SE*	*t*	*p*	95% CI	Status
H1	IM → AP	0.395	0.068	5.809	$< .001$	[0.262, 0.528]	Supported
H2	EM → AP	0.123	0.059	2.085	.040	[0.007, 0.239]	Supported
H3	IM → LA	0.618	0.051	12.118	$< .001$	[0.518, 0.718]	Supported
H4	EM → LA	0.244	0.073	3.342	.001	[0.101, 0.387]	Supported
H5	LA → AP	0.436	0.071	6.141	$< .001$	[0.297, 0.575]	Supported

Note. IM = Intrinsic Motivation; EM = Extrinsic Motivation; LA = Learner Autonomy; AP = Academic Performance; Bootstrapping = 5,000 subsamples.

All five direct-effect hypotheses were supported (Table 8). The IM → LA path was the largest coefficient in the model ($\beta = 0.618$), approximately 2.5 times the size of EM → LA ($\beta = 0.244$). Learner Autonomy was the strongest direct predictor of performance ($\beta = 0.436$), exceeding the direct effect of Intrinsic Motivation ($\beta = 0.395$). Extrinsic Motivation's direct contribution was modest ($\beta = 0.123$).

Figure 4

Structural Model Path Coefficients

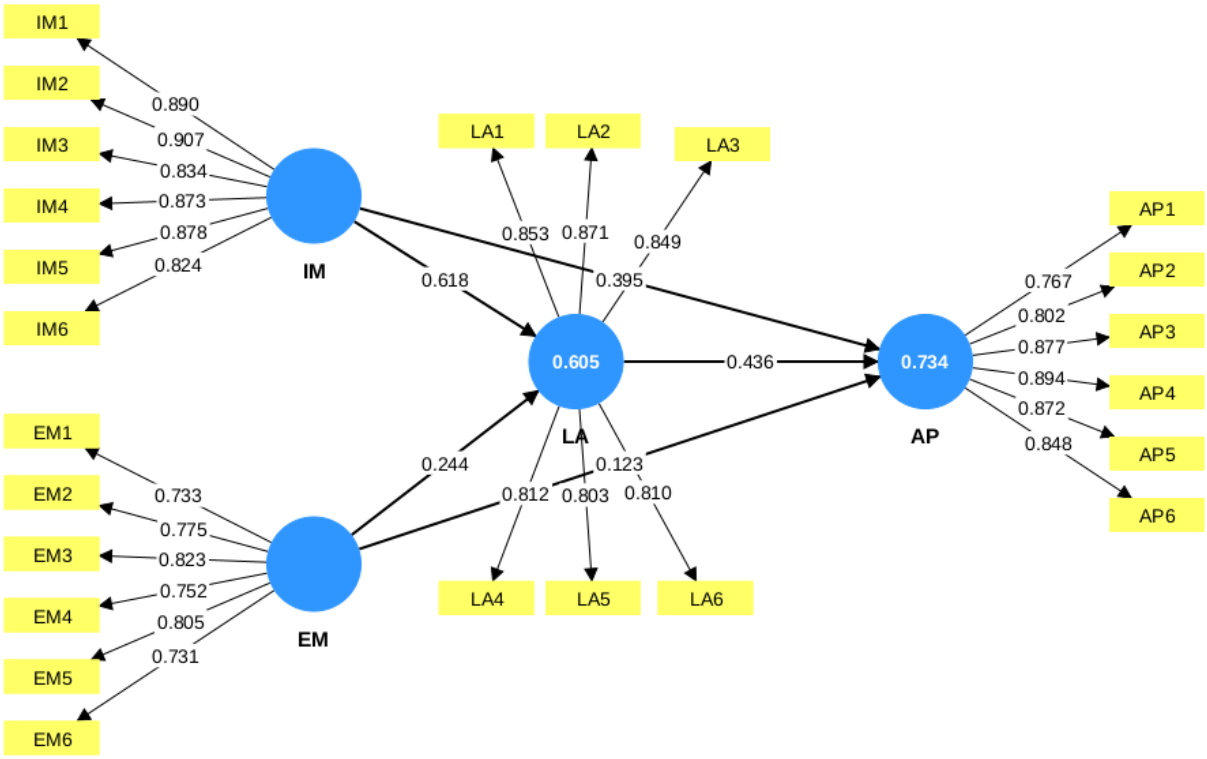


Figure 4 shows the standardized path coefficients and R^2 values for the endogenous constructs. The model explained 73.4% of the variance in Academic Performance ($R^2 = 0.734$; adjusted $R^2 = 0.728$) and 60.5% of the variance in Learner Autonomy ($R^2 = 0.605$; adjusted $R^2 = 0.600$).

Table 18

Effect Sizes (f^2) for Path Relationships

Path	f^2	Effect Size	Interpretation
IM → LA	0.684	Large	Substantial impact
LA → AP	0.282	Medium	Meaningful impact
IM → AP	0.245	Medium	Meaningful impact
EM → LA	0.106	Small	Limited impact
EM → AP	0.036	Small	Limited impact

Note. f^2 = Effect Size; thresholds: 0.02 = small, 0.15 = medium, 0.35 = large Cohen (1988).

The effect size for IM → LA ($f^2 = 0.684$) considerably exceeded Cohen's (1988) threshold for a large effect, and this effect size was approximately 6.5 times larger than that of EM → LA ($f^2 = 0.106$).

Table 19

R^2 Values for Endogenous Constructs

Construct	R^2	Adj. R^2	Predictors	Interpretation
Academic Performance	0.734	0.728	LA, IM, EM	Substantial

Construct	R^2	Adj. R^2	Predictors	Interpretation
Learner Autonomy	0.605	0.600	IM, EM	Moderate to Substantial

Note. R^2 values indicate proportion of variance explained; adjusted R^2 accounts for model complexity.

The combined predictors explained 72.8% of the variance in Academic Performance (adjusted $R^2 = 0.728$) and 60.0% of the variance in Learner Autonomy (adjusted $R^2 = 0.600$). Both values exceed benchmarks for substantial explanatory power in social science research Hair et al. (2019).

Table 20

Indirect Effects and Mediation Analysis

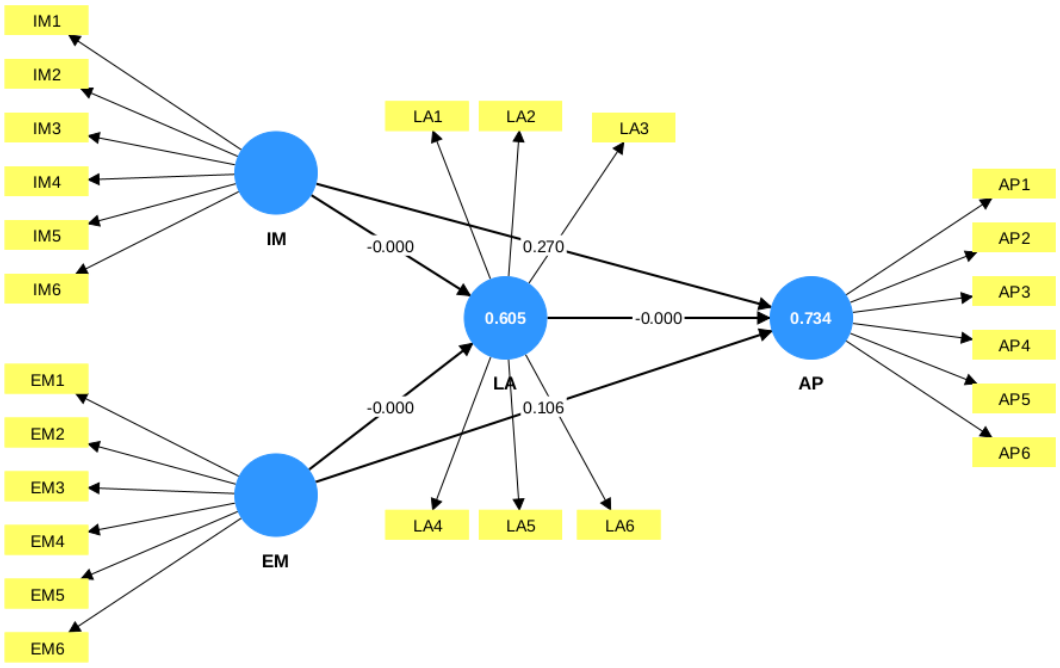
H	Mediation Path	Indirect β	*SE*	*t*	*p*	95% CI	VAF	Status
H6	IM → LA → AP	0.270	0.054	4.981	< .001	[0.163, 0.375]	40.5%	Supported
H7	EM → LA → AP	0.106	0.038	2.789	.006	[0.032, 0.188]	46.3%	Supported

Note. N = 145; bootstrapping = 5,000 subsamples; Mediation proportion = Indirect Effect / (Direct Effect + Indirect Effect).

Both mediation hypotheses were supported, indicating partial mediation. Learner Autonomy accounted for 40.5% of intrinsic motivation's total effect on performance (β_indirect = 0.270, p < .001) and 46.3% of extrinsic motivation's total effect (β_indirect = 0.106, p = .006). The higher proportion for extrinsic motivation indicates that external drivers depend more heavily on self-regulated learning behaviours to translate into academic outcomes.

Figure 5

Specific Indirect Effects through Learner Autonomy



Both indirect paths in Figure 5 were significant: intrinsic motivation through learner autonomy (β = 0.269, t = 4.981, p < .001) and extrinsic motivation through learner autonomy (β = 0.106, t = 2.789, p = .006). Although intrinsic motivation produced the larger absolute indirect effect, the mediated proportion was higher for extrinsic motivation (46.3% vs. 40.5%), indicating that external drivers depend more heavily on autonomous learning behaviours to affect performance.

4.2. Qualitative Findings

Thematic analysis of the twelve semi-structured interviews yielded four overarching themes, summarised in Table 12.

Table 21

Qualitative Themes and Key Findings

Theme	Core Finding	Supporting Hypotheses
Motivational Transition	Students evolve from external to internal motivation through internalization	H2, H4, H6, H7
Autonomy as Self-Regulation	Autonomy involves strategic help-seeking, metacognitive strategies, and resource curation	H3, H4, H5
Informal Immersion	Daily English media consumption bridges classroom and real-world proficiency	H1, H3, H6
Resilience Strategies	Students employ diverse coping mechanisms to recover from motivation fluctuations	H1, H2, H7

Theme 1: Motivational Transition.

The first theme concerns students' gradual movement from externally driven motivation to more internalized forms of engagement. Many participants reported that their initial reasons for learning English were strongly shaped by grades, family expectations, career plans, or the desire to meet academic requirements. However, by the third year, these external reasons had often become more personally meaningful. Rather than studying English only to satisfy external demands, students increasingly described learning as part of their own academic growth and future identity.

October clearly illustrated this shift: "In my first year, I mostly studied English because I cared about my GPA and did not want to fall behind. Later, I began to feel that improving my English was something I wanted for myself, not only something I had to do for school." Similarly, March explained, "Grades and teacher feedback still motivate me, but they are no longer the only reasons. I want to improve my English because I can see how it helps me become more confident and prepared for my future work." These accounts suggest that external motivators did not disappear; rather, they were gradually internalized and connected to students' personal goals.

Several participants also described motivation as a combination of internal interest and external discipline. For example, August stated, "My interest in English keeps me engaged, but external expectations also remind me to stay disciplined when I feel tired or distracted." This pattern is consistent with Self-Determination Theory, which argues that extrinsic motivation can support learning when external goals become personally valued. Overall, the theme shows that HUFLIT English-major students' motivation is not fixed but develops over time through the interaction between academic pressure, personal interest, and future-oriented goals.

Theme 2: Autonomy as Self-Regulation

The second theme reveals that learner autonomy among HUFLIT English majors was not a solitary or passive trait but an active repertoire of coordinated self-regulatory strategies. Students consistently described a structured approach to independent learning: they would first exhaust their own resources before seeking external help, and they deployed a range of metacognitive tools to monitor and direct their own progress.

Technology played a central role in this self-regulatory system. February, for instance, was deliberate in choosing which tools to engage: "I use lots of AI tools, but I don't use tools like ELSA, I use ChatGPT because it gives me lots of nuances in the words, in the word choice, word use, the meaning of the word, and also it diversified my structures." His use of artificial intelligence was not indiscriminate but tightly targeted at refining lexical precision and syntactic variety, a pattern that reflects the selective resource deployment characteristic of self-regulated learners. A different kind of resourcefulness emerged in November's strategy, which he called stage-skipping: "During that intermediate phase where I didn't know what to study, I deliberately sought out more advanced materials. By doing so, I could identify exactly what foundational knowledge I was lacking before tackling those advanced concepts. From there, I formulated a tailored roadmap to patch up those knowledge gaps." Rather than following a linear progression, he repurposed advanced content as a diagnostic instrument to locate his own weaknesses.

The sequential logic of independent work followed by strategic help-seeking appeared consistently across interviews. October captured the principle succinctly: "When encountering difficult vocabulary or confusing grammar structures, my first reflex is to play 'detective'. I look it up in dictionaries like Oxford or Cambridge, Google it, or search on language forums to solve the problem independently. If it takes too much time and I am still stuck, I will then ask my top-performing classmates or message the lecturer directly for help." This two-stage approach, in which learners maximized their own effort before turning to external sources, aligns closely with the forethought, performance, and self-reflection phases described in Zimmerman's (2002) cyclic model of self-regulated learning. The consistent pattern across participants indicates that autonomous learning, as practised by these students, was a deliberate and systematic process rather than an unstructured pursuit of independent study.

Theme 3: Informal Digital Immersion

The third theme captures how English had become embedded in students' daily lives through self-initiated exposure to digital media, functioning as a deliberate supplement to formal instruction. Participants described a range of informal immersion practices that they pursued independently of coursework, and they articulated clear reasons for separating these activities from academic obligations.

For some, informal media consumption doubled as both a learning objective and a motivational anchor. June explained, "One more thing that made me excited is that learning English will help me to understand movies, songs, and videos without subtitles." The prospect of unaided comprehension served as a tangible benchmark of progress, pulling her engagement forward. Others embedded English so thoroughly into daily routines that it ceased to feel like study at all. "I always search in English," January reported. "Try to expose yourself to English by searching in English. Try to do everything in English so you get used to it. You don't feel strange when you encounter new words or like you are in a scenery where you can only speak English." In her account, English had become a default operating medium rather than a subject to be mastered.

October went further, arguing that the motivational power of informal immersion depends on keeping it separate from academic demands: "I would advise you to spend 15-20 minutes every day consuming English content that you genuinely love -- whether that's watching vlogs, reading comics, or listening to podcasts. Just make sure it has absolutely nothing to do with your school textbooks. It will nurture your learning inspiration in the long run." This deliberate partitioning preserved the intrinsic enjoyment of the language and insulated it from the performance pressures of formal coursework. Across all accounts, informal immersion was consistently framed as a conscious pedagogical choice rather than passive entertainment. Students selected content, monitored their comprehension, and adjusted their practices over time, positioning these activities as an extension of, rather than a departure from, their formal language education.

Theme 4: Resilience Strategies

The fourth theme documents the diverse coping strategies that students employed to recover from periods of diminished motivation. Every participant reported experiencing motivational lapses, with common triggers including disappointing test results, heavy academic workloads, stalled progress, and competing personal demands. The strategies they described fell into four broad clusters: turning toward the language for relief, reaching forward through goal recall, responding structurally through planning, and turning outward toward social support.

The most vividly articulated strategy involved converting motivational setbacks into a matter of personal identity. July internalized this outlook into a phrase she repeated to herself: "I have a slogan to my mind at that time that you must learn, you must learn for yourself. If you don't eat, you cannot die. But if you don't study, you will die. So it becomes my go-to slogan." Framing learning as existential rather than optional transformed it from a task into a self-defining commitment. Others favoured a more pragmatic reset. May explained, "To overcome this, I usually take a short break then I make a simple study plan and start again." Her approach combined regulated rest with small, achievable targets that restored a perceived sense of agency over her circumstances.

Temporary disengagement from formal study was another recurring pattern, though its logic was counterintuitive: recovery came not from working harder but from stepping away. October recalled, "I put the theoretical textbooks aside for a few days. I found my inspiration again by watching English YouTube channels and podcasts about topics I actually enjoy, like psychology or movie reviews. It reminded me of my original reason for loving English, which gave me the momentum to get back on track with my university studies." By reconnecting with the intrinsic pleasure of the language through self-selected content, she replenished the motivational energy that academic routines had depleted. Across the sample, resilience did not appear to be a fixed trait but a set of strategies that students refined through repeated experience. Several participants noted that recovery from motivational lows became easier over time, suggesting that the capacity to bounce back from difficulty may consolidate into a durable disposition across successive cycles of challenge and recovery. The presence of multiple strategy types within individual participants further indicates that resilient students did not rely on a single coping mechanism but drew flexibly from a personal repertoire depending on the nature and severity of the motivational setback.

5. Discussion

The structural model accounted for 73.4% of the variance in academic performance ($R^2 = 0.734$; adjusted $R^2 = 0.728$) and 60.5% of the variance in learner autonomy ($R^2 = 0.605$; adjusted $R^2 = 0.600$), indicating that the motivation-autonomy-performance framework explains a substantial proportion of student outcomes in this context. This section synthesizes the quantitative and qualitative findings along three thematic lines, closing with an integrated assessment.

5.1. Intrinsic Motivation as the Primary Driver

As the results have shown, intrinsic motivation dominated the model by a considerable margin. Its effect on learner autonomy ($\beta = 0.618$, $t = 12.118$, $p < .001$, $f^2 = 0.684$) was more than two and a half times that of extrinsic motivation ($\beta = 0.244$, $t = 3.342$, $p = .001$, $f^2 = 0.106$). The direct path from intrinsic motivation to performance ($\beta = 0.395$, $t = 5.809$, $p < .001$, $f^2 = 0.245$) reached roughly three times the corresponding extrinsic route. This indicates that genuine interest in English functions as the primary driver of academic outcomes in this sample.

Interview data provide insight into the mechanisms underlying these coefficients. August stated that genuine interest in English eliminated the need for external prompting: “When you really like something, you don’t need anyone to tell you to do it. I automatically search for materials, watch videos, and practice without being assigned." This self-initiating quality is consistent with SDT’s claim that intrinsically motivated behaviour arises from its own satisfaction (Deci & Ryan, 2000). January described a similar mechanism in which long-standing exposure to English-language media established a perception of the language as accessible, generating a positive feedback loop between competence and continued engagement. October described a different trajectory, in which intrinsic motivation emerged not as an initial disposition but as an acquired orientation through the reorganization of her learning around authentic materials.

The disparity between the two coefficients (ratio exceeding 2.5 to 1) carries a practical implication. If the objective is to develop autonomous learners, instructional effort concentrated on cultivating genuine interest is likely to yield greater returns than effort directed at strengthening external incentives. The qualitative data suggest that interest cultivation is an environmental condition rather than a single intervention: participants who reported strong intrinsic motivation described exposure to authentic materials, instructors who modelled enthusiasm, and the gradual recognition that English could serve personal interests beyond the prescribed curriculum.

5.2. The Conditional Value of Extrinsic Motivation

Extrinsic motivation produced significant but substantively weaker effects. The EM → AP direct path reached significance ($\beta = 0.123$, $t = 2.085$, $p = .040$, $f^2 = 0.036$), as did the EM → LA path ($\beta = 0.244$, $t = 3.342$, $p = .001$, $f^2 = 0.106$); both were classified as small effects. The mediation results qualify this interpretation. Learner autonomy accounted for 46.3% of extrinsic motivation's total effect on academic performance, a higher proportion than the 40.5% observed for intrinsic motivation. The observation of higher proportion for extrinsic motivation reflects its smaller total effect, which translates to externally driven factors depend more heavily on the self-regulatory behaviours to produce measurable outcomes. When career requirements or institutional deadlines push students to set goals, monitor progress, and self-evaluate, those self-regulatory routines become the operative mechanism through which performance improves. The small direct effect of EM on AP further supports this interpretation;

absent the behavioural mediation of autonomy, extrinsic motivation carries limited independent force.

The qualitative evidence also points to variability in how external motivators function depending on their level of internalization. March offered a measured characterization by acknowledging that "grades and feedback are encouraging" yet emphasized that his primary drive was "to improve my English for myself, not just for school." This duality reflects SDT's internalization continuum in Figure 2. Students who remained nearer the controlled end of this continuum derived less benefit, while those who had achieved identified or integrated regulation, who could describe a career ambition as personally meaningful, were more likely to activate the autonomy pathway. This model aligns with Permzadian and Shen (2024), who found that expectancy-value constructs predicted outcomes largely through self-regulatory behaviour. In Vietnam, utility value served as the strongest motivational predictor because career-oriented goals, once internalized, drive sustained autonomous effort (Truong, 2025). A practical implication follows: where career pragmatism constitutes a central value, the evidence supports helping students connect external requirements to goals they personally value.

5.3. Learner Autonomy as the Key Mechanism

Among the three dependent variables, learner autonomy was the strongest direct predictor of academic performance ($\beta = 0.436$, $t = 6.141$, $p < .001$, $f^2 = 0.282$), surpassing the direct influence of intrinsic motivation. Autonomy functioned as the mechanism through which both forms of motivation channelled approximately 40 to 46 percent of their total effect, confirming the partial mediation hypothesized in **H6** and **H7**.

Qualitative data clarify the behaviours through which this mediation operates. October's systematic process of consulting dictionaries and online searches before approaching classmates or lecturers, February's use of ChatGPT to refine "nuances in word choice," and November's "stage-skipping" technique, which used advanced materials diagnostically, all exemplify the self-regulatory behaviours that Zimmerman's cyclic SRL model describes. These findings confirm that the forethought, performance, and self-reflection phases theorized as universal are observable among Vietnamese English majors. The data also reveal culturally inflected patterns: students frequently positioned autonomous behaviour as a compensatory response to instructional limitations, and informal immersion functioned as an extension of formal learning. This suggests that SRL in this context operates partly as a culturally adaptive mechanism enabling learners to bridge gaps between institutional resources and personal learning needs.

A tension in the autonomy data merits acknowledgement. The quantitative findings returned an LA mean of 3.97, the highest among all four constructs, yet several interview participants described significant discrepancies between their intentions and day-to-day execution (April's gaming distractions, September's lapses in focus, October's periods of feeling "overwhelmed, demotivated, and honestly, quite lazy"). These admissions do not invalidate the LA scores, as self-report instruments measure perceived autonomy rather than objectively observed behaviour. They do suggest that future research employing behavioural measures would strengthen claims about autonomous learning behaviour in this population.

5.4. Integrated Assessment

The structural model's explanatory power provides empirical support for the theoretical integration of SDT and SRL as a framework for understanding academic outcomes among

Vietnamese English majors. One contribution of the present study is the magnitude of the observed effects relative to benchmarks from Western EFL research. The large effect size for the IM → LA path ($f^2 = 0.684$) indicates that intrinsic motivation makes a particularly strong contribution to autonomy in this sample. This pattern is consistent with prior research linking self-determined motivation, self-regulation, and achievement (Noels et al., 2000; Zimmerman & Kitsantas, 2014). In a Vietnamese setting where teacher-centred instruction has historically constrained self-direction, strong intrinsic interest may function as a disproportionately powerful catalyst. SDT predicts that autonomy-supportive conditions amplify intrinsic engagement, and the present data are consistent with this prediction.

A further contribution concerns the mediation analysis. The finding that EM depends more heavily on the autonomy pathway (46.3%) than IM does (40.5%) refines SDT's internalization continuum by quantifying the differential dependence. Investing in intrinsic interest development strengthens performance through dual channels, direct cognitive engagement and indirect autonomy-driven learning. Relying predominantly on external incentives places heavier demand on the simultaneous cultivation of self-regulatory skills.

The qualitative data supply a developmental dimension that the structural coefficients alone cannot convey. The four interview themes outline a trajectory in which external pressures are progressively internalized, autonomous behaviours undergo iterative refinement, English becomes woven into daily life through self-selected media, and the capacity to recover from motivational setbacks consolidates into a durable disposition. This trajectory is not self-executing. The interviews indicate that it depends on exposure to effective teachers, access to authentic materials, opportunities to experience competence, and an institutional culture in which independent learning is treated as a baseline expectation. Where those enabling conditions are present, the statistical chain from motivation through autonomy to performance takes concrete form as a developmental pathway.

6. Conclusion

6.1. Summary of Findings

This study examined the interplay of motivation, learner autonomy, and academic performance among third-year English-major students at HUFLIT. Across the sample, students reported moderately high intrinsic motivation, a well-developed sense of ownership over their English learning, and broadly favourable perceptions of their academic progress. The structural model confirmed that the three predictors together captured the core motivational architecture of achievement in this population, yet beneath the uniformly positive self-reports a clear hierarchy of influence emerged.

Learner autonomy stood at the centre of the system. It was the strongest direct predictor of academic performance and served as the primary conduit through which both forms of motivation flowed toward achievement. Intrinsic motivation proved the more powerful of the two motivational sources, driving autonomy at more than double the strength of extrinsic motivation and exerting a meaningful direct effect on performance in its own right. Extrinsic motivation, while significant, depended proportionally more heavily on the autonomy pathway: although its absolute indirect effect was smaller than that of intrinsic motivation, a larger share of its total effect on performance was mediated through learner autonomy. In practical terms, this means that career ambitions and certification pressures at HUFLIT, where graduation requires IELTS 7.0 or TOEIC 780 alongside the Microsoft Office Specialist credential, translate into better grades only when students convert those external demands into concrete, self-regulated action.

The qualitative evidence showed that students did not simply study more; many had constructed systematic repertoires of autonomous practice that included setting specific learning goals, monitoring their own progress, selecting materials matched to their needs, integrating English-language media into daily routines, and deploying adaptive strategies to recover from motivational lulls. A recurrent gap between students' perceived autonomy and their day-to-day discipline suggested that HUFLIT English majors may overestimate their self-directed capacity, while individual lecturers surfaced repeatedly in interviews as pivotal figures whose influence on both motivation and autonomy fell outside the quantitative model. Taken together, the findings indicate that academic success for HUFLIT English majors depends not on being motivated or autonomous in isolation, but on building a self-reinforcing cycle in which genuine interest in the language fuels self-directed behaviour, and self-directed behaviour, in turn, deepens both competence and motivation over time.

6.2. Limitations

Several limitations constrain interpretation. The cross-sectional design prevents causal inference; data collected at a single time point cannot establish whether intrinsic motivation precedes autonomy development or whether the relationship is reciprocal. The reliance on self-report instruments introduces potential common-method bias, reflected in uniformly high means and the possibility of ceiling effects.

Academic performance was operationalised through self-reported perceived GPA and self-assessed proficiency rather than objective institutional records, a choice that reflects the theoretical commitments of the study. Within both SDT and SRL, self-evaluation and perceived competence function as motivational constructs in their own right rather than mere proxies for objective outcomes. An institutional GPA captures only final results, whereas perceived performance additionally encompasses the learner's sense of progress, self-appraisal, and academic self-concept, dimensions that are central to the frameworks guiding this research. The resulting limitation is therefore specific: our findings cannot speak to how motivation and autonomy relate to institutionally recorded grades. Future studies should address this by triangulating self-report measures with objective GPA data and standardized proficiency indicators.

6.3. Future Directions

Longitudinal designs tracking students across multiple semesters would permit examination of causal direction and developmental change. Incorporating objective performance measures (official GPA records, standardised test scores) and behavioural indicators would strengthen the evidentiary basis for claims about the autonomy-performance relationship. The qualitative findings regarding teacher influence suggest that future quantitative models should include instructor-related variables as potential moderators. The discipline gap between self-reported autonomy and acknowledged inconsistency warrants investigation through methods pairing self-report scales with time-use diaries or experience sampling. Comparative studies across multiple Vietnamese institutions and across Southeast Asian contexts would help clarify which patterns are institution-specific and which reflect more general features of motivation and autonomy development in collectivist educational cultures.

Acknowledgment

This study was developed from a student research project for the course Project 02: Scientific Research Methods. The research team sincerely thanks Dr. Tran Quang Bao Phuc for his guidance and feedback. The team also expresses appreciation to the Board of the Faculty

of Foreign Languages, HUFLIT, and the third-year English-major students who supported and participated in the study.

During the preparation of this manuscript, an LLM was used for language polishing and proofreading. The tool was used to improve phrasing, grammatical accuracy, and stylistic consistency. No AI tool generated research content, analyzed data, formulated interpretations, or produced the qualitative findings. All intellectual content, including the research design, data collection, statistical modelling, thematic analysis, and substantive argumentation, represents the original work of the authors. The authors reviewed all AI-assisted revisions and accept full responsibility for the accuracy and integrity of the final manuscript.

References

Afzal, H., Ali, I., Khan, M. A., & Hamid, K. (2010). A study of university students' motivation and its relationship with their academic performance. *International Journal of Business and Management*, *5*(4), 80. https://doi.org/10.5539/ijbm.v5n4p80

Alamer, A., Robat, E. S., Shirvan, M. E., & Ryan, R. (2025). Self-determination theory and language learning: A multilevel meta-analysis. *Educational Psychology Review*, *37*, Article 59. https://doi.org/10.1007/s10648-025-10038-y

Alhadabi, A., & Karpinski, A. C. (2020). Grit, self-efficacy, achievement orientation goals, and academic performance in university students. *International Journal of Adolescence and Youth*, *25*(1), 519–535. https://doi.org/10.1080/02673843.2019.1679202

Cao, D. T. P., & Pho, P. D. (2025). Relationship between learner autonomy and academic performance. *Issues in Language Studies*, *14*(2). https://doi.org/10.33736/ils.8441.2025

Ceylan, A. N. O. (2021). The relationship between learner autonomy and motivation. *The Turkish Online Journal of Educational Technology*, *20*(4). https://files.eric.ed.gov/fulltext/EJ1313495.pdf

Chong, S. W., & Reinders, H. (2022). Autonomy of English language learners: A scoping review of research and practice. *Language Teaching Research*, *29*(2), 607–632. https://doi.org/10.1177/13621688221075812

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Csizér, K., & Dörnyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *The Modern Language Journal*, *89*(1), 19–36. https://doi.org/10.1111/j.0026-7902.2005.00263.x

Cullen, S., & Oppenheimer, D. (2024). Choosing to learn: The importance of student autonomy in higher education. *Science Advances*, *10*(29), Article eado6759. https://doi.org/10.1126/sciadv.ado6759

Dang, T. T. (2010). Learner autonomy in EFL studies in Vietnam: A discussion from sociocultural perspective. *English Language Teaching*, *3*(2), 3. https://doi.org/10.5539/elt.v3n2p3

Danilenko, A., Kosmidis, I., Shershneva, V., & Vainshtein, Y. (2018). Learner autonomy in modern higher education. *SHS Web of Conferences*, *48*, Article 01022. https://doi.org/10.1051/shsconf/20184801022

Dao, V. H., Nguyen, M. D., & Pham, T. H. H. (2025). Developing self-learning skills in higher education in Vietnam: A survey of current practices and proposed solutions. *Nurture*, *19*(1), 38–53. https://doi.org/10.55951/nurture.v19i1.944

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/s15327965pli1104_01

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, *18*(3), 382. https://doi.org/10.2307/3150980

Garcia, T., & Pintrich, P. R. (1996). The effects of autonomy on motivation and performance in the college classroom. *Contemporary Educational Psychology*, *21*(4), 477–486. https://doi.org/10.1006/ceps.1996.0032

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). SAGE Publications.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hsieh, H.-C., & Hsieh, H.-L. (2019). Undergraduates' out-of-class learning: Exploring EFL students' autonomous learning behaviors and their usage of resources. *Education Sciences*, *9*(3), 159. https://doi.org/10.3390/educsci9030159

Ismail, S. M., Nikpoo, I., & Prasad, K. D. V. (2023). Promoting self-regulated learning, autonomy, and self-efficacy of EFL learners through authentic assessment in EFL classrooms. *Language Testing in Asia*, *13*(1), 27. https://doi.org/10.1186/s40468-023-00239-z

Minh, N. H., & Ngoc, N. T. D. (2023). Learner autonomy, motivation and English speaking proficiency: A study among English foreign language university students in Nghe An. *VNU Journal of Science: Education Research*, *39*(1). https://doi.org/10.25073/2588-1159/vnuer.4700

Nagle, C. (2021). Using expectancy value theory to understand motivation, persistence, and achievement in university-level foreign language learning. *Foreign Language Annals*, *54*(4), 1238–1256. https://doi.org/10.1111/flan.12569

Nakata, Y. (2014). Self-regulation: Why is it important for promoting learner autonomy in the school context? *Studies in Self-Access Learning Journal*, *5*(4), 342–356. https://doi.org/10.37237/050403

Nguyen, D. H. T. (2024). Vietnamese English majors' perspectives on teachers' motivational teaching at tertiary education. *PASAA*, *69*, 277–303. https://doi.org/10.58837/CHULA.PASAA.69.9

Nguyen, T. A., Nguyen, T. H., Nguyen, M. H., Phạm, T. T., Le, T. T., Phuong, Y. H., & Huynh, T. T. A. (2022). Empowering Vietnamese EFL learners at tertiary level: Investigating factors shaping learner autonomy in English language acquisition. *Vision: Journal for Language and Foreign Language Learning*, *11*(2), 125–138. https://doi.org/10.21580/vjv12i116406

Noels, K. A., Pelletier, L. G., Clement, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, *50*(1), 57–85. https://doi.org/10.1111/0023-8333.00111

Permzadian, V., & Shen, T. (2024). Assessing the predictive validity of expectancy theory for academic performance. *BMC Psychology*, *12*(1), 437. https://doi.org/10.1186/s40359-024-01935-y

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Merrill Prentice Hall.

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. https://doi.org/10.1037/0022-0663.98.1.209

Sinkkonen, M., & Tapani, A. (2024). Review of the concept "self-regulated learning": defined and used in different educational contexts. *International Journal on Social and Education Sciences*, *6*(1), 130–151. https://doi.org/10.46328/ijonses.640

Tao, S., & Jing, Y. (2023). More sense of self-discipline, less procrastination: The mediation of autonomous motivation. *Frontiers in Psychology*, *14*, Article 1268614. https://doi.org/10.3389/fpsyg.2023.1268614

Truong, B. C. (2025). The motivation of Vietnamese university students to learn English: A study using the expectancy-value model of academic motivation (Version 1). *Open Research Newcastle*. https://hdl.handle.net/1959.13/1333816

Truong, T. T. L., & Nguyen, V. L. (2024). Cultural influences on learner autonomy from the perspectives of Vietnamese EFL learners. *The Electronic Journal for English as a Second Language (TESL-EJ)*, *28*(2). https://doi.org/10.55593/ej.28110a5

Vallerand, R. J., & Bissonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality*, *60*(3), 599–620. https://doi.org/10.1111/j.1467-6494.1992.tb00922.x

Wang, Y., Luo, W., Liao, X., & Zhao, P. (2024). Exploring the effect of teacher autonomy support on Chinese EFL undergraduates' academic English speaking performance through the mediation of basic psychological needs and classroom engagement. *Frontiers in Psychology*, *15*, Article 1323713. https://doi.org/10.3389/fpsyg.2024.1323713

Wenden, A. L. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, *19*(4), 515–537. https://doi.org/10.1093/applin/19.4.515

Zhou, F., & Ahmad, R. A. H. R. (2025). The impact of learning motivation on students' academic performance: A self-determination theory perspective. *Journal of Education, Humanities, and Social Research*, *2*(1), 155–163. https://doi.org/10.71222/p1ck9h68

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*, *39*(2), 145–155. https://doi.org/10.1016/j.cedpsych.2014.03.004

Bionote

Ma Kinh Lac Che is an undergraduate student in the English Language programme at HUFLIT. His research interests center on the role of motivation in language learning and its measurement through structural equation modeling, particularly the application of quantitative methods to explore how intrinsic and extrinsic motivation shape academic outcomes in Vietnamese higher education.

Chau Hoang Anh Phan, also an undergraduate English Language student at HUFLIT, focuses on learner autonomy and self-regulated learning practices among tertiary EFL students. Her contribution to this study centers on understanding how students develop independent learning behaviours and how these behaviours translate into academic achievement.

Gia Bao Duong is an undergraduate English Language student at HUFLIT whose work combines quantitative and qualitative approaches. His research examines the lived experiences of English learners in Vietnamese university settings, using mixed methods to capture both the statistical structure and the subjective reality of language learning.

Minh Quyen Nguyen, an undergraduate English Language student at HUFLIT, is primarily interested in academic performance assessment and the psychological factors that influence English language achievement. His role in this study involved identifying the motivational and behavioural predictors that most strongly contribute to student success in EFL contexts.

.

#98 BEYOND THE PROMPT: A METACOGNITIVE ANALYSIS OF STUDENT AI INTEGRATION IN HIGHER EDUCATION

Hien Ngoc Nguyen Le[1]

Abstract

This study examines how third-year students at Vietnam Aviation Academy (VAA) utilize Generative Artificial Intelligence (GenAI) tools like ChatGPT and Google Gemini in their learning to enhance their linguistic capacity. As these tools become increasingly universal, there is a growing demand to understand how students in high-stakes English for Specific Purposes (ESP) programs can capitalize on Artificial Intelligence (AI)'s power to navigate their learning autonomy and maximize their language acquisition. Grounded in Flavell's (1979) metacognition theory, this study employed a convergent parallel mixed-methods design combining a structured survey of 62 students with their open-ended written responses. The current quantitative results highlighted a pronounced regulatory paradox: students planned AI use deliberately for lower-order tasks, characterized by vocabulary support and grammatical correction, yet they seldom verified AI outcomes against their course materials and adjusted their strategies for subsequent tasks. Reflexive thematic analysis based on four key themes (i.e, form-focused comfort zone, verification impasse, dependence-aware ambivalence, and scaffolding hunger) in which the last theme emerged inductively from data rather than the theoretical framework. AI integration offers two strands: students' cognitive offloading is not a knowledge deficit but an ecological adaptation to learning pressure as students use AI to encounter the threats of deadlines and academic competition. The paper concludes with pedagogical implications for integrating AI from an unstructured oracle into a disciplined cognitive mirror in ESP programmed at VAA.

Keywords: metacognitive regulation, generative artificial intelligence, cognitive offloading, aviation English, student perceptions.

1. Introduction

The proliferation of generative artificial intelligence (GenAI) has radically reshaped the realm of higher education and global language learning. As such, in the setting of English for Specific Purposes (ESP), the demand for maximizing the efficient acquisition of technical terminology and standardized communication is critical, and the integration of AI tools like ChatGPT and Gemini at the tertiary level has become indispensable. Existing research confirms that AI can accelerate dynamic learning ecosystems where students can exploit instant feedback and personalize their learning paths (Alshammari & Alhamazany, 2026; Olaide et al., 2024), and these benefits drive high adoption and form strong trust among second language (L2) learners (Kılıçkaya & Kic-Drgas, 2025).

Despite the remarkable benefits of AI adaptation in education, AI poses a pedagogical concern when it erodes students' learning autonomy and cognitive competence (Misiejuk et al., 2026). Current research warns that unstructured AI use can trigger cognitive offloading and metacognitive disengagement (K. Fan & Gao, 2026; Vendrell & Johnston, 2026; Wang et al., 2025), noting that over-reliance on AI-generated output can replace students' self-monitoring skills. Similarly, Qu et al. (2025) conducted a meta-analysis grounded on Bloom's taxonomy, indicating that although GenAI has positive effects on lower-order outcomes with vocabulary

[1] *Vietnam Aviation Academy, Ho Chi Minh City, Vietnam, Email: hienlnn@vaa.edu.vn*

memorization and grammatical exercises, it has limitations in fostering higher-order skills such as analysis, evaluation, and critical thinking when students' scaffolding is low. In the same vein, Vendrell and Johnston (2026) indicate a prevalent phenomenon of fluency heuristic when students conflate their AI-generated outcome with their true mastery: students tend to uncritically accept AI's output and ignore the necessary process of reflective trial-and-error to form their understanding ((Tomisu et al., 2025). Notably, although AI is beneficial in consolidating learners' learning trust, it can impair their performance due to metacognitive overload (McCarthy et al., 2018)–a state when students feel overwhelmed with managing AI prompts and hallucinations, driving them to withdraw effortful critical thinking and lean toward passive cognitive offloading (Jingjing et al., 2026). The pedagogical challenge, therefore, is nothing about whether AI is allowed in class, yet how AI regulations can be cultivated without triggering either blind trust or exhausting students' necessary vigilance for hallucination prevention.

These tensions are crucial for VAA students, as they still face a severe intrinsic cognitive load: while managing uneven entry-level English proficiency, they must master the same amount of technical aviation terminology within demanding coursework. It was also observed that under unguided conditions, students may use Gen AI extensively without proper cognitive evaluation, leading to an inability to detect context-specific errors or algorithmic hallucinations and placing them on the brink of cognitive offloading. While a growing number of international studies have investigated student perceptions of GenAI use largely in Western contexts such as the United States, Germany, and Finland (Ketzer-Nöltge & Rüger, 2025; McCarthy et al., 2018; Misiejuk et al., 2026), no study has examined how ESP students in developing countries in Southeast Asia, for example, Vietnam, adapt strategic self-regulation in dealing with cognitive offloading caused by extensive AI integration in learning. This constitutes a combined contextual and empirical gap for this study. To address this gap, the present study employed a convergent parallel mixed methods design to examine the metacognitive awareness, AI usage behavior, and self-regulated learning patterns of VAA students. The study was set to third-year aviation English students at a single institution during the second semester and the study is guided by three research questions:

RQ1. To what extent do third-year VAA students demonstrate metacognitive awareness of generative AI tools in learning aviation English?

RQ2. What are their prevailing usage configurations and attitudes toward generative AI tools in learning aviation English?

RQ3. How do these students perceive and navigate the metacognitive tension between AI-driven task efficiency and the potential loss of autonomous language-learning skills?

By generating localized empirical evidence, this study proposes concrete, evidence-based principles for educators to scaffold metacognitive AI adoption into disciplined cognitive reflection, transitioning students from surface-level query and paste behaviors to deep self-regulated learning.

Scope and delimitations. This study is anchored in four core respects. First, for participant cohort, the study prioritizes on third-year students of aviation English majors at Vietnam Aviation Academy (VAA)- as this group carrying the heaviest simultaneous load of general and safety-critical technical English- compared to other cohorts of students or disciplines. Second, as for educational tool, the study concentrates on assistance of generative AI in students' daily academic use, specifically with ChatGPT and Google Gemini though they are not licensed completely at the level of universities. Third, regarding timeframe: the survey

was administered on 7 and 8 April 2026, at the end of a semester that ran from late December 2025 to early April 2026, a cross-sectional window that allows the researcher to capture learning habits. Lastly, for measurement, the study is based on self-report about how students perceive and use AI in their everyday study, hence capturing their metacognitive awareness and regulation of AI use rather than objectively observing their behavioral error tracking of AI.

2. Literature Review

Extensive research on AI empowered education shows that AI offers both opportunities and challenges in the setting of teaching and learning at tertiary level, concurrently holding the potential to foster students' metacognitive, self-regulated learning. It is noted that without careful instruction of AI use, AI can pose severe pedagogical threats due to extensive reliance on AI's output, weakening critical thinking among students. To comprehensively unpack these complexities, the current literature is organized by three primary subthemes: (1) AI affordance in ESP learning, (2) metacognitive awareness in AI-assisted learning, (3) the risks of unstructured use of AI, (4) ambivalent engagement of AI

2.1. AI affordances in ESP learning

A number of studies have been undertaken to explore why L2 learners leverage GenAI's power to maximize their learning paths, in which AI chatbots are frequently mentioned as cognitive and motivational scaffold (Dahri et al., 2024; Jingjing et al., 2026). Functionally, Ai tools are designed to deliver instant, personalized feedback, they are great resources to mitigate students' anxiety and adapt explanation to their capacity accordingly (Alshammari & Alhamazany, 2026; Karmazinuk & Helik, 2026). In collaborative work, chatbots operates as conversational partners that explicitly foster affective reinforcement to producing language (Kılıçkaya & Kic-Drgas, 2025). Among benefits, the aspect of trust formation is highlighted when the AI integration rate is increasing with long-term research, demonstrating that institutional support and perceived utility cultivate a trong intention to use AI for self-regulated learning (Dahri et al., 2024). In the realm of ESP learning, AI affordances offer relief from defining burdens- processing complex technical texts, navigating complex academic materials- operating as a psychological safety net that mitigates technostress and prevent cognitive overload (Misiejuk et al., 2026; Shen & Tao, 2025) however, adoption and satisfaction are measured far more frequently in this research tradition rather than what learners actually use AI, which is where metacognition becomes theoretically and practically valuable.

2.2. Metacognitive awareness in AI-assisted learning

In the context of AI-mediated language learning, metacognitive awareness has emerged as a key referee to position whether generative AI serves as a learning scaffold or a cognitive crutch. Empirically, scholars concur that effective metacognitive regulation, which encompasses three main processes from planning to monitoring and finally evaluating projects, can directly reduce students' writing anxiety but significantly enhance their self-efficacy (Shen & Tao, 2025). Teng (2025) also demonstrated that when EFL learners are cultivated with metacognitive awareness, they can proactively plan, monitor, and evaluate ChatGPT feedback productively, thereby managing their cognitive load more effectively and developing more diverse, adaptive task strategies. In other words, high-metacognitive learners generally utilize AI as a "cognitive mirror" to deepen their reflections (Tomisu et al., 2025) and externalize their thought processes (Wang et al., 2025), whereas students with lower metacognitive awareness are prone to "fluency heuristic," in which GenAI leads students to mistake surface-level polish for factual accuracy and abandon critical self-reflection (Vendrell & Johnston, 2026). Consequently, these students are in favor of "metacognitive laziness" (Y. Fan et al., 2025) and

passive cognitive offloading. Under this circumstance, Rogti and Ouarniki (2026) suggested AI-powered scaffolding can strengthen students' language acquisition only when explicitly trained. This recognition matters for the present study as ESP learners, driven by heavy academic workload, need to know how to proactively plan AI usage not to maximize immediate tasks without critical monitoring while preserving their deep, autonomous learning

2.3. The risks of unstructured use of AI

A wealth of research highlighted that unstructured GenAI use generates high cognitive costs; thus, this study identifies three primary AI-generated risks: cognitive offloading, fluency heuristic, and cognitive overload. First of all, the most significant of these risks is known as "cognitive offloading." Raitskaya and Tikhonova (2025) caution that AI can trigger the phenomenon of cognitive offloading—that is, when students uncritically accept AI-generated ideas and neglect necessary evaluation of authenticity and quality (Misiejuk et al., 2026; Qu et al., 2025); heavy reliance on this act is associated with lower critical thinking performance (Gerlich, 2025) . Consequently, students frequently fall into a state of "illusion of mastery" where they conflate the AI's proficiency with their genuine capacities, drifting students away from learning autonomy (Raitskaya & Tikhonova, 2025). Critically, scholars concur with AI risks, yet diverge on its mechanism. For instance, Vendrell and Johnston (2026) determine another risk related to the fluency heuristic: when students are flavor in surface-level polish of AI-generated output over factual accuracy, they tend to skip verification stages prior to submission. Tomisu et al. (2025) identify another risk of unstructured AI use in lost trial-and-error when AI as oracle routinely supplies the correct solutions, making students lose their drive to observe and adjust their mistakes. This elimination of desirable difficulties cultivates a notion that Fan et al., (2025) describe as a passive state of "metacognitive laziness". Conversely, scholars including Jingjing et al., (2026), and McCarthy et al.,(2018) mentioned the different sets of risks: demanding that students frequently monitor AI outputs results in metacognitive overload and technostress. According to Fan and Gao (2026), these threats are conditional, and AI-mediated support can be detrimental for learners when used without proper guidance. Despite the convergence on risks, researchers have mostly investigated risks of unstructured configuration of AI use separately, while in general cohorts it is unclear how such AI -generated risks compound in ESP settings where undetectable terminology errors are safety-relevant.

2.4 Ambivalent engagement of AI

The growing body of research consistently reports positive effects of AI such as mitigating students' linguistic anxiety and building self-efficacy, for most of which GenAi is usually mentioned as a psychological safety net (Shen & Tao, 2025; Zhai & Nezakatgoo, 2025). This consensus, however, captures only a half of learners' experience perceptions. Ketzer-Nöltge and Rüger (2025) hold that the very students who used to praise AI for its usefulness (i.e., saving time and streamlining tasks) turn out to voice serious concerns (fearing its risks) about their overdependence and skill loss within academic integrity, diminishing independent thought. Indeed, students cannot identify states of "loving or hating AI" as they experience an affective paradox where they feel both simultaneously. This reflects two distinct dimensions of AI integration in education, yet they do not contradict each other. Beyond this, it is crucial to understand a central question unexamined: whether worries about AI discipline students' self-regulation and learning behaviors when enthusiasm and worry coexist and how students perceive and navigate their own fear of losing autonomous language learning skills.

Despite the increasing number of studies investigating learners' metacognition and AI adoption, significant gaps still remain. It is observed that most studies are conducted in Western

contexts such as Finland, Germany, and the United States (Ketzer-Nöltge & Rüger, 2025; McCarthy et al., 2018; Misiejuk et al., 2026), yet none investigates how these mechanisms can shape the metacognitive configuration for ESP learners in developing Southeast Asian countries, specifically in Vietnam's ESP programs. Under the intense curricular pressure, third-year students majoring in aviation at VAA frequently carry a heavy learning load caused by dense terminology and uneven entry-level proficiency, driving students to tensions between AI adoption for lower-order efficiency and metacognitive offloading. Furthermore, existing empirical studies have relied on single-strand quantitative designs, specifically SEM-based surveys (K. Fan & Gao, 2026; Shen & Tao, 2025), analysis of logs (Misiejuk et al., 2026) and experiments (McCarthy et al., 2018; Sun et al., 2025) while there are some notable mixed methods research from Dahri et al. (2024), Teng, (2025a), (2025b) and Wang et al. (2025) who conducted their studies in EFL cohorts outside Southeast Asia, leaving no study that has mapped awareness, usage configuration, perceptions, and concerns about AI in Vietnam. Beyond this, this study employs a convergent mixed-methods approach that allows the author to integrate statistical analysis with students' own written accounts, thereby providing a more cohesive understanding of how Vietnamese ESP students navigate AI integration. This study also addresses the need for offering contextual and cultural insights by investigating four variables—learners' knowledge and regulation of their own cognition that is formulated by theoretical framework of Flavell (1979).

3. Theoretical Framework: Metacognition in AI-Assisted Learning

This study is grounded in Flavell's (1979) theory of metacognition, which describes how learners understand and control their own thinking through two interconnected components: metacognitive knowledge – what learners know about persons, tasks and strategies- and metacognitive regulation- which encompasses the active monitoring and evaluation of their learning progress during and after a task. Originally, Flavell (1979) introduced metacognition to explain how learners "know about knowing", the concept of metacognition has since been integrated into models of self-regulated learning, evolving a standard analytical lens in EFL and educational technology research (Teng, 2025a; Wang et al., 2025). The introduction of generative AI fundamentally reshapes both components: metacognitive knowledge and metacognitive regulation. For the construct of metacognitive knowledge, under the AI integration in learning, it now encompasses an accurate understanding of what AI can and cannot perform, involving its linguistic fluency, hallucinations, and blind spots related to domain-specific content. Regarding metacognitive regulation, it also requires learners to proactively verify machine outputs compared with their own developing knowledge- which is a rigorous demand that the previous study skills do not insist. Notably, when knowledge is present without self-regulation, the learners may recognize AI-generated errors without procedures to detect these errors- which is also the precise configuration that this study investigates.

Figure 1 illustrates the study's conceptual framework which is anchored by the theory of metacognition of Flavell (1979). This theory grounds four measured constructs: metacognitive awareness of AI- shaping both usage configurations and attitudes that contributes to metacognitive tensions between valued efficacy and feared autonomous loss. In turn, tensions can help define the pedagogical scaffolding for educators (see Fig. 1).

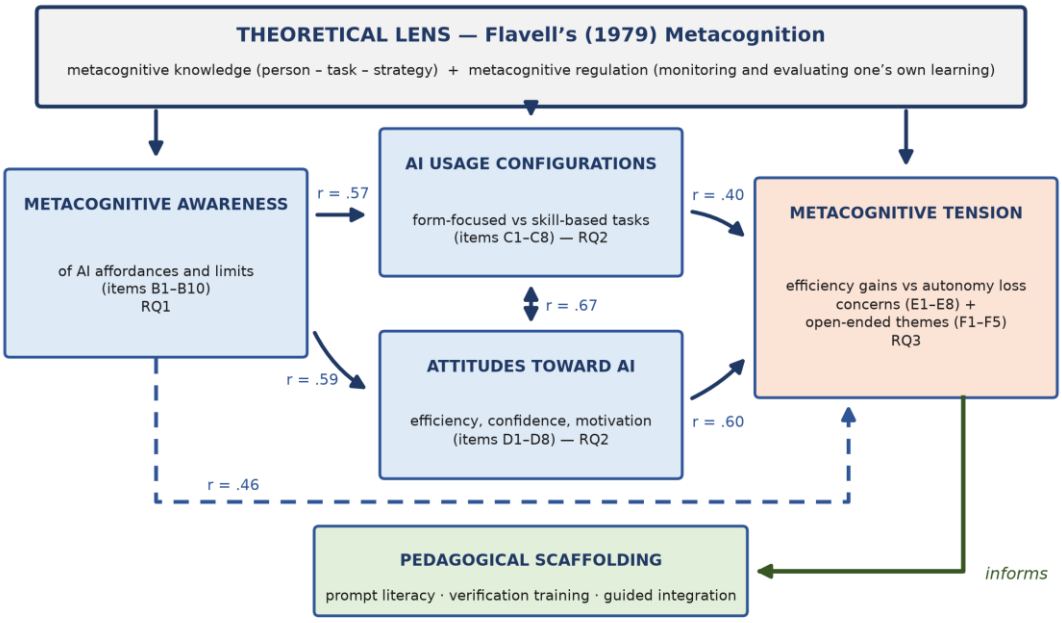


Figure 1. Conceptual framework: Flavell's (1979) metacognition applied to AI-assisted aviation English learning, with the study's four constructs, research questions, and observed inter-construct correlations (all $p < .01$).

Metacognition is considered as an effective framework for this study, which conceptualizes on four grounds that corresponds to the criteria of explanatory match, contextual relevance and scholarly credibility. First, regarding explanatory fit: the phenomenon under investigation- which is centric on students who frequently delegate cognitive work to AI while consciously being aware of consequences like deteriorating critical thinking ability for tasks- is at its core an issue of self-regulation. Meanwhile, strategic self-regulation is the primary mediator to address AI-triggered cognitive offloading in L2 learning (Misiejuk et al., 2026; Wang et al., 2025). Second, this framework aligns with conceptual requirement with two components connect directly into the study's measures. As for metacognitive knowledge, it is operationalized as awareness of AI's affordances and limits while metacognitive regulation is captured in stages ranging from verification, adjustment, to appraisal behaviors which are assessed by usage, attitude, and concern scales. For the branch of knowledge regulation division, it is set to measure the core analytical distinction between declarative awareness and evaluative competence (Teng, 2025a). Third, the paradigm has a contextual relevance as the theory concerns internal cognitive processes not just focusing on culturally specific behaviors, hence its applications come from Asian EFL settings comparable to Vietnam (Shen & Tao, 2025; Teng, 2025b; Wang et al., 2025), and it extends to a domain of ESP where verifying one's own understanding is safety relevant. Fourth, as for scholarly credibility, it is acknowledged that cognitive load theory explains the mental cost of dense aviation English that students must carry, and technology acceptation models (TAM) can elaborate on technological adoption in education (Dahri et al., 2024), yet it is not in focus. Self-determination theory illuminate motivation ((Zhai & Nezakatgoo, 2025), but not mentioning the knowledge -verification stage at the heart of this study. Therefore, cognitive load, acceptance and motivation serve as contextual condition while metacognition carries analytical work. Driven by this model, the research questions examine students' metacognitive

knowledge of AI, the usage configurations and attitudes towards how AI is enacted, and the tension that may occur when regulation falters under efficiency pressure.

4. Methodology

This study was part of larger mixed-methods research conducted at an institution, Vietnam Aviation Academy (VAA), in Ho Chi Minh City during the second semester of the academic year, specifically from late December 2025 to April 2026. At that time, GenAI became an everyday tool across Vietnamese higher education. By that time, tools like ChatGPT and Google Gemini became embedded in how students read, wrote, and practiced English, yet there was not any formal instruction or pedagogical guidance on when and how far students rely on them. Although AI technology constantly evolves, the questions related to this strategy—what students know about their AI use in their daily learning, how they configure it for their study, and how they weigh its benefits against risks—still remain both timely and significant to explore. This research focused on third-year students, aviation English majors, as they carry a heavy language load due to dense terminology. While newer data may reveal additional insights to the field, the patterns observed from this study remain relevant today. By focusing on this critical timeframe, the current dataset allows a cohesive analysis of students' metacognition in the conditions where the reliance on AI use is high or low.

4.1. Research Design

The study adopted a convergent mixed methods design (Creswell & Creswell, 2018) that was implemented though a single bilingual (English-Vietnamese) instrument to collect quantitative and qualitative data in parallel. The instrument contains 34 Likert-scale items measuring the four constructs in Figure 1, paired with five open-ended questions eliciting students' own perceptions about autonomous AI use, efficiency gains, challenges, higher-order effects and curricular suggestions for educators. The quantitative strand was set to address the first and second research questions; meanwhile the open-ended questions helped to deepen the third research questions, and the two strands are merged as a cohesive interpretation to evaluate convergence.

4.2. Research Context and Participants

The current study was conducted among third-year students at Vietnam Aviation Academy (VAA), ho Chi Minh City, during the second semester of the academic year in the aviation English program. These students indeed undergo double load of learning: developing general English while maintaining safety-critical aviation terminology and communication protocols. The survey was distributed to the cohort via Google Forms and collected 80 submissions, yet after data cleaning, only 62 unique responses were retained. Most respondent were aged from 18 to 23 (M = 19.5, SD = 1.3; six did not report a usable age); the male number was 30, leaving 27 as females, and 5 preferred not to mention. As for sef-assessed proficiency, candidates were clustered at B1–B2 (n = 32) and A1–A2 (n = 26), with 4 students at C1–C2; for AI habitual use, 41 of 62 students reported using AI for their study at least 3 times to 4 times per week (at 66%), while the remaining 21 used it one to two times per week. From the survey, it indicated that ChatGPT dominated at 82% while Google Gemini accounted for 79%, followed by AI Google Translation (45%), Grammarly (19%), Microsoft Copilot (16%), DeepL (10%), and Quill Bot (6%). Table 1 summarizes the sample's characteristics.

Table 1

Participant Characteristics (N = 62)

Characteristic	**n**	**%**
Gender		
Male	30	48
Female	27	44
Prefer not to say	5	8
English proficiency (self-assessed)		
Beginner (A1–A2)	26	42
Intermediate (B1–B2)	32	52
Advanced (C1–C2)	4	6
Frequency of AI use for English		
1–2 times per week	21	34
3–4 times per week	16	26
5–6 times per week	8	13
Daily (7+ times per week)	17	27

Note. N = 62. Age M = 19.5, SD = 1.3 (range 18–23); six respondents did not report a usable age. Percentages are rounded and may not total 100.

4.3. Instrument

The questionnaire comprised six primary sections: section A collected demographic and usage frequency of AI in their study, while section B measures metacognitive awareness of AI with 10 items (e.g., "I can identify when AI-generated English content contains errors or inappropriate language") that had adapted from operationalization's of metacognitive awareness in AI-assisted EFL learning studies from Teng (2025) and Wang et al., 2025). As for section C, it was set to measure students' usage configurations across eight tasks types, including vocabulary, grammar support to speaking activities and aviation-specific simultaneous test such as drafting Air Traffic Control (ATC) communications, translating technical safety manuals, and developing scripts for cabin crew role-play scenarios and video projects. Section D with 8 items were to gauge student attitudes toward AI functionalities – which are usefulness, confidence, motivation for AI collaboration- while section E focused on AI-triggered- concerns with 8 items ranging from accuracy, dependence, autonomous loss, ethics. Systematically, all items from section B to E were based on five-point Likert scales (agreement scales for sections B, D and E, and a frequency scale from never to very often for section C), whereas section F devoted for the open-ended questions where students could present both in Vietnamese and English (see Appendix A). the bilingual format was a deliberate choice as it ensured that students with limited constraint of English could express their own thinking deeply and accurately.

4.4. Data Cleaning

The questionnaire was distributed to the full third-year cohort from 2 classes via Google Forms, and the export contained 80 rows. Every row was screened before analysis. Eighteen

rows repeated the submission timestamp of an earlier row to the millisecond, appeared out of chronological order at the end of the export and differed from the earlier row in 0 to 8 fields. They were treated as repeated submissions and removed, keeping the earlier row of each pair, and no values were changed. An earlier version of this paper retained 17 of these rows and excluded one unique submission, reporting N = 78; all analyses in this version were re-run on the 62 unique submissions. Six free-text age entries could not be interpreted and were treated as missing. The wording of awareness item B5 was overwritten in the data export and was restored from the original questionnaire. All Likert responses were converted into their numeric values ranging from 1 to 5 for data analysis.

4.5. Data Analysis

Quantitative analysis followed three steps rigorously. Firstly, the researcher assessed the internal consistency of each scale by using Cronbach's alpha and corrected item-total correlations (Table 2). Subsequently, the researcher profiled each construct and item with their descriptive statistics before checking their inferential contrasts: paired samples t-tests compared declarative with evaluative awareness and form-focused with skill-based usage. Pearson correlations were used to examine the relationships among the four core constructs. For open-ended responses, they were subjected to analyses through reflective thematic analysis, following Braun and Clarke (2012) six-phase framework and reflecting a constructivist stance. Codes and themes accordingly were developed inductively from data; for example, with Vietnamese's responses, the research was analyzed in the original and translated only for reporting to keep their true nuances of interpretation, and, in line with the reflexive approach, themes are reported without prevalence counts; 55 respondents gave at least one substantive written response. Because many responses were brief, the themes are exploratory and are used mainly to explain and illustrate the quantitative results. At the final stage, the two strands were combined through side-by-side comparison of each quantitative construct with its corresponding qualitative theme (Creswell & Creswell, 2018), the phase that convergences and divergence between numbers and the narratives become visible.

4.6 Trustworthiness and ethics

The findings of four scales indicated that they are highly reliable. As shown in Table 2, all alphas exceeded 0.90 and every corrected item and total correlation exceeded 0.71, meaning that there was excellent. The item set was suitable for factor analysis, Kaiser–Meyer–Olkin = .84 and Bartlett's test of sphericity $\chi^2(561) = 3010.8$, $p < .001$. A principal-axis extraction returned a single dominant factor per scale on which all standardized loadings were high (.72 to .97), and the average variance extracted (AVE = .67–.86) and composite reliability (CR = .94–.98) exceeded the conventional .50 and .70 thresholds, confirming convergent validity. Discriminant validity was supported by the Fornell–Larcker criterion: the square root of each construct's AVE (.82–.92) exceeded that construct's correlations with the other three constructs (Table 5). However, all items were worded in the same direction and many students chose the same answer across items, so these high values may partly reflect uniform responding rather than construct quality alone, and the validity indices and correlations should be read with caution.

Table 2

Reliability and Validity of the Four Scales (N = 62)

Scale (construct)	Items	α	Item–total range	Loading range	AVE	CR
B. Metacognitive awareness	10	.97	.84–.92	.85–.93	.80	.98
C. Usage configurations	8	.94	.75–.85	.77–.89	.67	.94
D. Attitudes	8	.98	.88–.95	.89–.97	.86	.98
E. Concerns	8	.95	.71–.89	.72–.91	.72	.95

Note. Alphas computed on the cleaned sample after removal of duplicate submissions.

For the qualitative strand, the researcher maintained and audit trail linking of every theme to its coded excerpts and rigorously checked all translations of Vietnamese's quotations against the originals. Participants, who were voluntary and anonymous, were offered a consent form to clearly explain the purpose of this research, and submission of the questionnaire also constitutes consent. No identifying information was collected, and all quotations are reported without identifiers.

5. Results

This section devotes for presenting the results in accordance with three main research questions that are categorized into 5 subsections: section 1 focusing on metacognitive awareness of research question 1 (RQ1), section 2 for identifying usage configurations and attitudes of VAA students, (RQ2) section 3 addressing quantitatively concerns associated with student attitudes (RQ3), section 4 presenting qualitative themes and section 5 integrating the strands together.

5.1. RQ1: Metacognitive Awareness

First, the research looked at how VAA students perceived what AI chatbots could and could not do to facilitate them in study (RQ1). From this, it indicated that although VAA students' awareness of AI benefits and perils were at moderate level on average (M=3.54, SD=1.16), yet the ten-item trajectory revealed more information than the mean, in which students concurred the most for understood the strengths and imitations of AI-generated content (B2, M=3.74). whereas they least agreed with the two items that require applying that knowledge—being able to spot errors in AI output (B8, M=3.37) and considering AI suggestions critically rather than just accepting them (B10, M=3.37). To test this difference directly, a comparison was made between each student's average score on the three declarative items (B1-B3, M=3.60) with their average score on the three evaluative items (B4,B8,B10, M=3.44), $t(61) = 2.08$, $p = .042$, $dz = 0.26$. ultimately, these findings reveal a critical disconnect: although VAA students recognize AI's fallibility, they lack the evaluative confidence to detect AI-generated errors in practice.

Table 3

Construct and Selected Item Descriptives (N = 62)

Construct / item	M	SD
B. Metacognitive awareness (10 items)	**3.54**	**1.16**

B1 Understand what AI can and cannot do	3.48	1.41
B2 Aware of strengths and limitations of AI content	3.74	1.30
B3 Plan the specific help needed before using AI	3.58	1.33
B4 Can evaluate whether AI output is accurate/appropriate	3.58	1.19
B5 Monitor own understanding when using AI	3.63	1.26
B6 Reflect on how AI affects learning progress	3.58	1.27
B7 Know when to rely on AI vs. other resources	3.60	1.30
B8 Can identify errors in AI-generated content	3.37	1.22
B9 Adjust AI use to what works best	3.50	1.29
B10 Think critically rather than accept AI suggestions	3.37	1.27
C. Usage configurations (8 items)	**3.48**	**0.92**
C1 Vocabulary support	3.58	1.11
C2 Writing revision	3.66	0.97
C3 Grammar checking and correction	3.69	1.03
C4 Translation (Vietnamese–English)	3.55	1.15
C5 Reading comprehension support	3.58	1.06
C6 Pronunciation practice and feedback	3.26	1.12
C7 Speaking practice	3.08	1.21
C8 Aviation-specific terminology and protocols	3.42	1.05
D. Attitudes (8 items)	**3.75**	**1.11**
D1 Positive attitude toward AI for English	3.81	1.17
D2 AI useful for improving English skills	3.76	1.17
D3 AI makes learning more efficient / saves time	3.79	1.15
D4 AI increases confidence in using English	3.61	1.18
D5 AI increases motivation to learn English	3.76	1.18
D6 AI helps achieve better grades	3.82	1.17
D7 AI provides personalized support	3.79	1.19
D8 Would recommend AI to other students	3.68	1.28
E. Concerns (8 items)	**3.68**	**0.95**
E1 Concern about accuracy of AI content	3.60	1.06
E2 Worry AI use may be cheating / dishonesty	3.69	1.06
E3 Fear of becoming too dependent on AI	3.82	1.03
E4 AI may prevent developing autonomous skills	3.58	1.11

E5 Worry AI weakens critical thinking	3.68	1.05
E6 Concern about privacy and data security	3.69	1.15
E7 Uncertain about ethical implications	3.68	1.10
E8 Worry about culturally inappropriate output	3.66	1.17

Note. All items scored 1–5. Full item wording appears in Appendix A.

5.2. RQ2: Usage Configurations and Attitudes

The study subsequently proceeded to investigate students' AI reconfiguration and their AI toward them (RQ2). Although majority of students (66%) admitted that they frequently used AI for their English study at least 3 times per week, their use of AI is not uniform across tasks. For instance, the rates were varied for form-focused tasks including grammar checking (M = 3.69), writing revision (M = 3.66), vocabulary support (M = 3.58), and translation (M = 3.55). Notably, as for skill focused group: pronunciation, aviation-specific terminology, and speaking practice gained the least common use with M = 3.26, M = 3.42, and M = 3.08 respectively, VAA students fairly use AI for reading support (M=3.58). A paired samples t test comparing the composite means reveals that students scored significantly higher on form-focused tasks, ranging from C1 to C4 with M=3.62 than on skill-based tasks, ranging from C5 to C8 with M= 3.33. This means a moderate effect size ($t(61) = 3.84$, $p < .001$, $dz = 0.49$; see Table 4), hence this discrepancy illustrates a paradoxical pedagogical challenge within the aviation English curriculum. Specifically, students at least likely use AI to practice speaking although they recognize the significance of oral proficiency in their professional field.

Table 4

Paired-Samples t Test: Form-Focused versus Skill-Based AI Usage (N = 62)

Comparison	**M_1 (SD)**	**M_2 (SD)**	**ΔM [95% CI]**	**t(61)**	**p**	**dz**
Form-focused (C1–C4) vs. skill-based (C5–C8)	3.62 (0.97)	3.33 (0.96)	0.29 [0.14, 0.44]	3.84	< .001	0.49

Note. M_1 = form-focused configurations (C1–C4); M_2 = skill-based configurations (C5–C8). ΔM = mean difference; dz = Cohen's dz for paired samples. Items scored 1–5.

Out of the for constructs, attitudes towards AI were the most favorable (M = 3.75, SD = 1.11) as students believed that AI can enhance their learning motivation (D5, M = 3.76), while helping them gain better grades (D6, M = 3.82), and providing support that is suited to each individual needs (D7, M = 3.79). The two strands- AI usage and student attitudes- have a strong association and they are consistent with the study's conceptual framework (Figure 1), in which awareness, use, and attitude reinforce one another.

Table 5

Pearson Correlations Among the Four Constructs (N = 62)

Construct	**1**	**2**	**3**	**4**
1. Metacognitive awareness	—			

2. Usage configurations	.57	—		
3. Attitudes	.59	.67	—	
4. Concerns	.46	.40	.60	—

Note. All correlations are statistically significant at $p < .01$. Construct scores are respondent means across scale items (1–5).

5.3. RQ3: The Tension Between Efficiency and Autonomy

Concerns about AI were almost as high as attitudes (M=3.68, SD=0.95). Specifically, the fear that students endorsed most strongly was the feelings of becoming overly dependent on AI and losing the ability to learn on their own (E3, M = 3.82), which was observed as one of the highest-rated items in the survey, level with the most positive attitude items. Worries about weakened critical thinking (E5, M=3.68), reduced autonomous learning (E4, M=3.58), and the accuracy of AI output (E1, M=3.60) followed. The key result for RQ3 is that attitudes and concerns rose together rather than trading off against each other: the two were positively correlated ($r = .60$, $p < .001$), meaning the students who valued AI most were often the same students who worried about it most. In fact, 52% of respondents scored 3.5 or higher on both attitudes and concerns at once. The cohort is therefore not divided into enthusiasts and skeptics; most students hold both feelings together, experiencing the tensions between efficiency and autonomy within themselves. The size of this correlation partly reflects uniform responding, since it fell to $r = .25$ among the 22 students who varied their answers across the questionnaire; the more robust observation is that most students were high on both scales.

Table 6

Cross-Tabulation of Attitudes and Concerns (N = 62)

	Concerns high (≥ 3.5)	**Concerns low (< 3.5)**	**Total**
Attitudes high (≥ 3.5)	32	9	41
Attitudes low (< 3.5)	8	13	21
Total	40	22	62

Note. Cell entries are counts of respondents. The 32 students high on both constructs represent 52% of the sample—the "ambivalent engagers" described in the text.

5.4. RQ3 (Qualitative): Four Interconnected Themes

To understand these patterns from the students' own point of view, we analyzed their answers to the five open-ended questions (55 students gave substantive written responses). Four themes emerged, and together they tell a connected story: students settle into a comfortable routine of form-focused use ***(Theme 1)***; when they suspect the AI's outputs are wrong, and they have no way to check them ***(Theme 2);*** students lack the ability to verify errors, leaving them openly ambivalent about becoming dependent on AI ***(Theme 3);*** and they request the institution to teach them how to use it well ***(Theme 4).*** Table 7 summarizes the themes, their sub-themes, and the quantitative construct each one helps explain.

Table 7

Thematic Map of Open-Ended Responses (n = 55 substantive respondents)

Theme	Sub-themes	Linked construct
1. Form-focused comfort zone	on-demand lexico-grammatical repair; low-anxiety conversational rehearsal	C (usage); D (efficiency, confidence)
2. Verification impasse	suspicion without a checking procedure; prompt-communication breakdowns	B (evaluative gap); E1 (accuracy)
3. Dependence-aware ambivalence	gratitude–worry coexistence; conditional "if abused" framing	D × E correlation; E3 (dependence)
4. Scaffolding hunger	requests for prompt training and rules; aviation role-play proposals	Pedagogical implications

Note. Themes were developed through reflexive thematic analysis and are reported without prevalence counts. Quotations are verbatim; those marked (VN) are translated from Vietnamese.

Theme 1: *The form-focused comfort zone.* Qualitative responses revealed that students overwhelmingly rely on AI to repair language forms (for instance, translating, checking grammar, and editing essays) as AI can outperform it, and especially without feeling judged. One student mentioned this typical habit: "outside class, i use AI to talk about my daily life in English, then use AI to transform my idea to a more academic version" On the same vein, another highlighted AI's efficiency as a time-saver, saying "Thay vì mất hàng giờ tìm kiếm tài liệu, AI có thể tóm tắt nội dung bài đọc, tạo danh sách từ vựng từ một video YouTube hoặc soạn nhanh các bài tập" ["Instead of spending hours searching for materials, AI can summarise a reading, make a vocabulary list from a YouTube video or quickly prepare exercises"] (VN) Additionally, a notable subgroup of respondents utilized AI to practice speaking in a low-pressure environment, and deliberately regarding AI as a peer for human interaction. It is true that while AI offer "comfort zone" for students by providing fast, personalized and low-anxiety support to students, they paradoxically raise a significant pedagogical concern. Such fear is largely about the usage patterns that students constructed with AI, heavily lowering cognitive abilities where, previous research also emphasizing, the transformative educational benefits of AI are least impactful.

Theme 2: *The verification impasse*. Regarding their challenges, the majority of students claimed that they experienced doubt without a viable method for resolution, hence accepting AI to err and even leading them to wrong answer without noticing. While they may suspect that AI was incorrect, they lacked strategies to verify its output effectively. One respondent described this uncertainty directly: "Một khó khăn là AI không phải lúc nào cũng chính xác 100%, nên đôi khi tôi không chắc câu trả lời có đúng hay không" ["One difficulty is that AI is not always 100% accurate, so sometimes I am not sure whether the answer is right or not"] (VN). Another wrote, "One challenge I face is that AI sometimes gives incorrect or unclear answers. I also feel that I may rely too much on AI instead of thinking by myself. Therefore, I need to check the information carefully and use AI wisely," yet students kept confirming that they remained unable to articulate exactly how conduct this verification. In these open-ended responses, a small subset of participants traced these difficulties back to their own prompting skills, noting "tôi gặp khó khăn trong việc sử dụng các câu lệnh để AI có thể hiểu rõ được vấn đề mà tôi đang gặp phải" ["I have difficulty using prompts so that AI can clearly understand

the problem I am facing"] (VN). This theme is to represent the qualitative manifestation of the evaluative gap in the qualitative survey, specifically shedding more light on the question items B8 and B10, as they received the lowest awareness scores, indicating that precise evaluative weakness in their own words.

Theme 3: *Dependence-aware ambivalence*. As for the higher-order impacts of AI in student's metacognition, VAA candidates acknowledged that AI can extend their ideas if used as a support tool, although AI can weaken their cognitive capabilities if students are overly reliant on it. One student wrote, "Over-reliance: i might become too lazy and stop thinking for myself if i use AI too much." Another explained this mechanism: "AI cung cấp nhiều góc nhìn khác nhau cho một vấn đề, giúp học cách so sánh, đối chiếu và đánh giá. Tuy nhiên, nếu lạm dụng quá có thể mất đi khả năng tự suy luận, dễ trở nên lười biếng" ["AI offers many perspectives on a problem, which helps with learning to compare, contrast and evaluate. However, overusing it can take away the ability to reason for oneself and easily make one lazy"] (VN). Interestingly, many students framed this risk conditionally by stating that "if students are too lazy to think for themselves and let AI do all the brainstorming, they are using AI in the wrong way" in the open-ended responses, students may project the danger onto a hypothetical careless student. Finally, this theme acts as the qualitive embodiment of the attitude-concern correlation in the survey since enthusiasm and worry are not held by different students but voiced by the same ones.

Theme 4: *Scaffolding hunger.* The most practical finding that the study raised is that student fail to advocate for reduced AI integration; rather, they explicitly request formal instruction from educators and institutions on how to utilize AI more ethically and responsibly. Their recommendations clustered around three primary requests: training in prompt formulation, structured boundaries for AI use and AI allowance in aviation learning context. Regarding prompt trainings, students suggested "Nên học cách prompt cho AI tối ưu trong việc hỗ trợ học tập" ["We should learn how to prompt AI optimally to support studying"] (VN), while students with a need for structure boundaries also claimed "AI should be integrated as a guided support tool in a blended learning model, with clear rules and training, so students can use it effectively without becoming dependent." Third, when it comes to AI application for ESP in teaching and learning, one member insisted "Aviation role-play: use AI to simulate realistic situations, like conversations bewteen [sic] pilots and ATC." Therefore, students intuitively articulated a demand for properly guided AI integration in higher education as a coherent pedagogical instruction.

5.5. Integration of Strands

Collectively, the two main strands converge in the same direction at every step of framework. Specifically, the gap between form-focused and skill-based use ($dz = 0.49$) is depicted in Theme 1's comfort zone, while the evaluative awareness gap- scored the lowest- is presented in Theme 2's verification impasse. Next, Theme 3's ambivalence devotes for the mutual relationship between attitudes and concerns, leaving the last theme- Theme 4 for naming the remedy—training in prompting and verification, clear rules, and aviation-specific practice—which turns the study's diagnosis into its practical recommendations.

6. Discussion

The overarching objective of this research was to investigate the ways how third-year students majored in aviation English truly comprehend, practically apply and navigate their concerns regarding GenAI in their learning path. The synthesized results reveal a notable paradox: students form solid declarative knowledge, limit their reliance on AI primarily to

form-focused mechanism, and articulate profound concerns, whereas they fundamentally lack the evaluative strategies and they request institutional frameworks to convert this awareness into regulated, ethical integration. There are four critical insights representing as follow.

6.1. Awareness Is Declarative, Not Yet Evaluative

The moderate awareness scores conceal a significant matter that students recognize that AI may make mistakes due to its limitations, yet they cannot detect those limits in specific outputs. This adds necessary nuance to Teng's (2025) optimistic view of metacognitive awareness in AI-assisted writing, demonstrating that although students can acknowledge AI's flaws, it does not guarantee that students can have the operational ability to verify its output or whether AI hallucinates. Compared with Flavell's (1979) terms, VAA students possess metacognitive knowledge without the corresponding regulation, which further explains that this dynamic directly mirrors the "expertise paradox" and the challenge of "trust calibration" identified by Stockwell and Wang (2025). Therefore, to properly establish trust in AI, a user must possess the expertise that AI is used to supplement and facilitate them rather than lead them. As students are still acquiring language proficiency, they undeniably lack linguistic expertise to independently evaluate the AI's hallucination, leading them to trust AI system blindly under uncertainties (Stockwell & Wang, 2025). In the settings of ESP, this distinction is not merely academic as an unverified term can cause a real-world hazard. When students described "I don't know how to verify the answers"- indicating verification impasses- it identifies the exact evaluative skills that Vietnam Aviation Academy must target.

.*6.2. The Comfort Zone Confirms the Lower-Order Bias—With an Aviation Irony*

Rather skill-based configurations, students have significant preferences for form-focused replicates, at the level of everyday use, the meta-analytic finding reveals that AI can be reliable at lower-order outcomes (Qu et al., 2025). Naturally, students will gravitate toward grammar repair and translation tasks as these applications provide an immediate, verifiable feelings and anxiety free environment (Zhai & Nezakatgoo, 2025). This perfectly aligns with recent findings that AI chatbots can serve as a non-judgmental linguistic safety net to help mitigate language apprehension for students. Students, though preferring comfort within this lower-order zone, may risk sever cognitive offloading, since they habitually bypass the deeper thinking engagement for advanced communication. This situation, therefore, creates a profound irony that while speaking is known as the exact skill for their aviation processional, students in this cohort treat it with the least skills practiced with AI. Furthermore, the students who adopted AI for conversational practice demonstrate that the configuration is technologically friendly, yet AI is simply not pedagogical scaffolded into the majority's repertoire, ultimately leaving them trapped in dilemma of mechanical fixed rather than authentic conversational rehearsal.

6.3. Ambivalent Engagement, Not Adoption Versus Resistance (RQ3)

The strongest and least expectation from this quantitative finding is the positive correlation between attitudes and concerns ($r = .60$), with half scoring high on both aspects. The acceptance literature tends to show the opposite poles of enthusiasm and worry (Dahri et al., 2024); such data illustrate the both factors - enthusiasm and worry- are rising together, echoing Ketzer-Nöltge and Rüger's (2025) observation that students are holding sophisticated metacognitions regarding AI's costs though appreciating its benefits. As for the fear of over dependance on AI, the highest rated concern among the survey's top items overall serves not as brake on students' AI use but as a running commentary along it. Theme 3 helps show how they manage to use AI psychologically: some feel guilty when overusing AI for most of their

tasks, but they have to do it due to heavy workload survival. To manage their sense of guilt over relying on AI, students use conditional framing- as they identify that AI is only harmful if abused- projecting the risk of laziness for careless students rather than themselves. In their open-ended responses, VAA students recognized that the dense terminology makes it too demanding for them to verify AI output , and they are short of practical strategies to fix this issue. Hence, students, although aware of AI's hallucination, lack necessary verification skills to audit it, unable to gain better study habits. Beside this unresolved tensions exhausts students' metal resources, hance leaving students falling into technostress and metacognitive overload (Jingjing et al., 2026; McCarthy et al., 2018).

6.4. Implementation

This study offers implication extracted from theme 4. First, for prompt and verification skill, it is crucial to set up short, explicit training sessions for students to teach them how to construct prompts and critically cross check AI output against course materials, characterizing glossaries, and authoritative aviation references, converting the verification impasse into a routine. Then, guided configuration shift is a must for students. Specifically, this study suggests that instructors must design well structures tasks that AI use up the skill ladder that students frequently avoid – the most obvious application for this is to establish AI-simulated pilot–controller dialogues – which also align with the cognitive mirror principle of transitioning AI as a partner that elicits rather than hinder their own thinking (Tomisu et al., 2025; Vendrell & Johnston, 2026). Finally, it is recommended that clear institutional rules should be applied to establish explicit usage guidelines, legitimizing AI as a supportive tool rather than strictly constraining it in education. The current absence of such legislation feeds both unguided misuse and digital anxiety. By contrast, clear institutional scaffolding is like a leverage to ensure that AI tool acceptance translates into strategic self-regulation (K. Fan & Gao, 2026), effectively buffering students against language learning apprehension and technostress (Shen & Tao, 2025). Since these measures directly impact the survival needs that students articulated, adoption resistance is eradicated naturally.

6.5. Limitations and Future Research

While the research offers critical insight into student's metacognition development with AI integration, its limitations must be acknowledged. The sample comes from only one institution with one cohort of third year aviation English students (N=62), so generalization beyond English for Specific Purpose (ESP) requires caution. All quantitative measures are self-reported, and evaluative competence may need behavioral testing to gauge error detection tasks with seeded hallucinations. The study employed cross-sectional design, unable to track how ambivalence resolves- whether it turns into regulation or into resignation- over time. The qualitative strand consists of writing open-ended responses, which are broader but shallower than interviews, while follow-up interviews could probe the conditional framing – identified in Theme 3. Finally, the high alpha alphas (ranging from 0.94 to .98) may indicate some item redundancy within scales, hence future versions of the instrument could be shortened to ensure information fulfilment. Finally, longitudinal designs paring usage logs with proficiency outcomes would identify whether the form-focused comfort zone can retard higher-order growth like what the offloading literature predicts.

7. Conclusion

This study was anchored on Flavell's (1979) metacognition theory, combined with mixed methods to map how third-year students at VAA know, use and form concerns about GenAI in their study path. The findings compose a single coherent picture: students have

awareness of strengths and weaknesses of GenAI, and their use is habitual, focusing on a form-focused comfort zone. Besides, VAA students have positive attitudes toward AI chatbots due to being shadowed by fear of dependence. For the open-ended responses, they helped trace the mainstream mechanism related to certain results: unable to verify AI output, students heavily rely on AI's suggestions, this acting on their own misgivings, and their well-articulated ambivalence. Geographically, the study contributed to empirical evidence in an under researched Southeast Asian, specifically in the realm of English for Specific Purpose (ESP), and aviation specific practice. The current study highlighted that the barriers to healthy AI incorporation in education are not derived from students' ignorance or resistance but from an unmet institutional guidance to evaluate how to use them more properly. Programs that allow scaffolding the needs can transform an unstructured oracle into a disciplined cognitive mirror before comfort zone of AI turns into dependence.

Acknowledgments

The author expresses sincere gratitude to the administration and third-year aviation English students at the Vietnam Aviation Academy for their invaluable participation and support. Special thanks are owed to Dr. Thu Le for her expert instructional guidance, which significantly shaped the methodological rigor of this study.

References

Alshammari, S. H., & Alhamazany, A. (2026). Artificial Intelligence-Supported English as a Foreign Language Learning: A Narrative Review of Personalisation, Feedback, Engagement and Pedagogical Collaboration. *International Journal of Technology in Education and Science*, *10*(2), 219–236. https://doi.org/10.46328/ijtes.7018

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology: Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon*, *10*(8), e29317. https://doi.org/10.1016/j.heliyon.2024.e29317

Fan, K., & Gao, T. (2026). The role of AI-mediated metacognitive interventions in EFL writing development: A cross-sectional study of textual accuracy, lexical sophistication, and cohesion. *Acta Psychologica*, *267*, 107069. https://doi.org/10.1016/j.actpsy.2026.107069

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, *15*(1), 6. https://doi.org/10.3390/soc15010006

Jingjing, Q., Yan, Z., Xiaobing, X., Li, L., & Nusrat, A. (2026). The impact of technostress and generative AI on EFL mathematic metacognitive reading strategies and digital classroom burnout in SDG-aligned education. *Scientific Reports*, *16*(1), 17026. https://doi.org/10.1038/s41598-026-47683-4

Karmazinuk, K., & Helik, J. (2026). Bridging technical and emotional skill gaps: AI-enhanced adaptive learning and emotional intelligence in project management education. *Research in Learning Technology*, *34*. https://doi.org/10.25304/rlt.v34.3625

Ketzer-Nöltge, A., & Rüger, A. (2025). Meta-cognition on AI: What students think about using AI for academic purposes. *Cognition, Communication, Discourse*, *30*, 53–80. https://doi.org/10.26565/2218-2926-2025-30-04

Kılıçkaya, F., & Kic-Drgas, J. (2025). Enhancing Collaborative Learning Practices via Chatbots: Insights from an EFL context. *SAGE Open*, *15*(3), 21582440251369189. https://doi.org/10.1177/21582440251369189

McCarthy, K. S., Likens, A. D., Johnson, A. M., Guerrero, T. A., & McNamara, D. S. (2018). Metacognitive Overload!: Positive and Negative Effects of Metacognitive Prompts in an Intelligent Tutoring System. *International Journal of Artificial Intelligence in Education*, *28*(3), 420–438. https://doi.org/10.1007/s40593-018-0164-5

Misiejuk, K., López-Pernas, S., Kaliisa, R., & Saqr, M. (2026). Cognitive offloading in student–AI collaboration: A longitudinal analysis of prompting strategies. *Computers in Human Behavior Reports*, *22*, 101130. https://doi.org/10.1016/j.chbr.2026.101130

Olaide, A. A., Nathaniel, S., Oyesola, O. A., & Abayomi, A. K. (2024). Artificial intelligence with agile teaching strategies for personalized learning. *Nigerian Online Journal of Educational Sciences and Technology (NOJEST)*, *6*(1). https://nojest.unilag.edu.ng/article/view/2139

Qu, X., Sherwood, J., Liu, P., & Aleisa, N. (2025). Generative AI Tools in Higher Education: A Meta-Analysis of Cognitive Impact. *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–9. https://doi.org/10.1145/3706599.3719841

Raitskaya, L., & Tikhonova, E. (2025). Enhancing Critical Thinking Skills in ChatGPT-Human Interaction: A Scoping Review. *Journal of Language and Education*, *11*(2), 5–19. https://doi.org/10.17323/jle.2025.27387

Rogti, M., & Ouarniki, O. (2026). From error correction to strategy support: AI-powered metacognitive scaffolding in assessing EFL writing performance and self-regulatory engagement. *Language Teaching Research*, 13621688261420086. https://doi.org/10.1177/13621688261420086

Shen, X., & Tao, Y. (2025). Metacognitive Strategies, AI-based Writing Self-efficacy and Writing Anxiety in AI-assisted Writing Contexts: A Structural Equation Modeling Analysis. *International Journal of TESOL Studies*, *7*(1), 70–87. https://doi.org/10.58304/ijts.20250105

Stockwell, G., & Wang, Y. (2025). Framing human-AI collaboration in language education. *Technology in Language Teaching & Learning*, *7*(4), 103824. https://doi.org/10.29140/tltl.v7n4.103824

Sun, X., Dou, W., & Yang, Y. (2025). The socio-emotional dangers of using Artificial Intelligence (AI) technologies in second language (L2) education: Unveiling Chinese EFL teachers' perceptions and experiences. *Acta Psychologica*, *261*, 105956. https://doi.org/10.1016/j.actpsy.2025.105956

Teng, M. F. (2025a). Metacognitive Awareness and EFL Learners' Perceptions and Experiences in Utilising CHATGPT for Writing Feedback. *European Journal of Education*, *60*(1), e12811. https://doi.org/10.1111/ejed.12811

Teng, M. F. (2025b). Understanding EFL student writers' metacognitive awareness in utilizing ChatGPT. *System*, *135*, 103848. https://doi.org/10.1016/j.system.2025.103848

Tomisu, H., Ueda, J., & Yamanaka, T. (2025). The cognitive mirror: A framework for AI-powered metacognition and self-regulated learning. *Frontiers in Education*, *10*, 1697554. https://doi.org/10.3389/feduc.2025.1697554

Vendrell, M., & Johnston, S.-K. (2026). Scaffolding critical thinking with generative AI: Design principles for integrating large language models in higher education. *Computers*

and Education: Artificial Intelligence, *10*, 100572. https://doi.org/10.1016/j.caeai.2026.100572

Wang, P., Liu, T., Yang, Y., & Xiang, X. (2025). Optimizing self-regulated learning: A mixed-methods study on GAI's impact on undergraduate task strategies and metacognition. *British Journal of Educational Technology*, bjet.70018. https://doi.org/10.1111/bjet.70018

Zhai, Y., & Nezakatgoo, B. (2025). Evaluating AI-Powered Applications for Enhancing Undergraduate Students' Metacognitive Strategies, Self-Determined Motivation, and Social Learning in English Language Education. *Scientific Reports*, *15*(1), 35199. https://doi.org/10.1038/s41598-025-19118-z

Bionote

Hien Ngoc Nguyen Le is an English language education lecturer at the Vietnam Aviation Academy. With extensive teaching experience across prominent Vietnamese higher-education institutions, her research interests focus on educational technology integration, specifically the impact of AI-powered assistance on metacognitive regulation and self-regulated learning in English for Specific Purposes contexts.

Appendices

Appendix A: Survey Instrument

The instrument was administered bilingually (English and Vietnamese); only the English stems are reproduced here. Items in Sections B, D and E used a five-point agreement scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree), and items in Section C used a five-point frequency scale (1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = very often). No items were reverse-keyed. Participation was voluntary and anonymous; informed consent was obtained on the opening screen.

A.1. Section A — Participant information

A1. Age (years).

A2. Gender: Male / Female / Prefer not to say / Other.

A3. How would you rate your current English proficiency level? Beginner (A1–A2) / Intermediate (B1–B2) / Advanced (C1–C2).

A4. How often do you use AI tools (e.g., ChatGPT, Gemini) for learning English per week? 1–2 times / 3–4 times / 5–6 times / Daily (7+ times).

A5. Which AI tools do you use for learning English? (check all that apply)

A.2. Section B — Metacognitive awareness of AI (10 items)

B1. I understand what AI tools can and cannot do when helping me learn English.

B2. I am aware of the strengths and limitations of AI-generated content for language learning.

B3. Before using AI tools, I think about what specific help I need (e.g., vocabulary, grammar, writing feedback).

B4. I can evaluate whether the English language output from AI tools is accurate and appropriate.

B5. I actively monitor my understanding when using AI tools to learn new English concepts.

B6. I reflect on how AI assistance affects my English learning progress.

B7. I know when to rely on AI tools and when to use other learning resources (e.g., teachers, textbooks, peers).

B8. I can identify when AI-generated English content contains errors or inappropriate language.

B9. I adjust my use of AI tools based on what works best for my English learning goals.

B10. I think critically about AI suggestions rather than accepting them without question.

A.3. Section C — AI usage by task type (8 items)

Stem: "How often do you use AI tools for the following activities?"

C1. Vocabulary support (e.g., finding word meanings, synonyms, example sentences).

C2. Writing revision (e.g., improving essays, reports, or assignments).

C3. Grammar checking and correction.

C4. Translation between Vietnamese and English.

C5. Reading comprehension support (e.g., summarizing texts, explaining difficult passages).

C6. Pronunciation practice and feedback.

C7. Speaking practice (e.g., conversation simulation, dialogue generation).

C8. Learning aviation-specific English terminology and communication protocols.

A.4. Section D — Attitudes toward AI (8 items)

D1. Overall, I have a positive attitude toward using AI tools for learning English.

D2. AI tools are useful for improving my English language skills.

D3. Using AI tools makes learning English more efficient and saves me time.

D4. AI tools help me feel more confident in using English.

D5. AI tools increase my motivation to learn English.

D6. I believe AI tools can help me achieve better grades in English courses.

D7. AI tools provide personalized learning support that meets my individual needs.

D8. I would recommend AI tools to other students learning English.

A.5. Section E — Concerns about AI (8 items)

E1. I am concerned about the accuracy of English language content generated by AI tools.

E2. I worry that using AI tools for assignments might be considered cheating or academic dishonesty.

E3. I am concerned that I might become too dependent on AI tools and lose my ability to learn independently.

E4. Using AI tools might prevent me from developing important autonomous learning skills.

E5. I am worried that relying on AI tools will weaken my critical thinking skills in English.

E6. I have concerns about the privacy and security of my data when using AI tools.

E7. I am uncertain about the ethical implications of using AI for academic work.

E8. I worry that AI tools might provide culturally inappropriate or contextually incorrect English language suggestions.

The five open-ended questions in Section F are reproduced in Appendix B.

Appendix B: Open-Ended Questions

The five open-ended questions followed the Likert sections of the questionnaire. Each question was presented in English and Vietnamese, and students could answer in either language. The written responses were analysed through reflexive thematic analysis (Section 4.5) and form the qualitative strand reported in Section 5.4.

F1. How do you typically use AI tools for autonomous (self-directed) English learning outside of class? Please describe specific examples.

F2. In what ways have AI tools helped you learn English more efficiently? What specific time savings or learning gains have you experienced?

F3. What challenges or difficulties have you encountered when using AI tools for learning English?

F4. How do you think using AI tools affects your development of higher-order language skills (e.g., critical thinking, creative writing, complex communication, analytical reading)?

F5. What suggestions do you have for how AI tools should be integrated into the English curriculum at Vietnam Aviation Academy? What guidelines or support would help students use AI more effectively?

#99 AI-POWERED GAME-BASED VOCABULARY PRACTICE FOR ENHANCING VOCABULARY RETENTION AMONG YOUNG VIETNAMESE EFL LEARNERS: A QUASI-EXPERIMENTAL STUDY

Trang Trinh Khanh Lien[1], Chi Ho Phuong Nguyen[2]

Abstract

This quasi-experimental study examined whether Blooket-supported game-based vocabulary practice enhanced vocabulary retention among young Vietnamese EFL learners. Eighty-two A2-level learners aged 10–11 at a private language centre in Ho Chi Minh City participated in an eight-week intervention. The experimental group (n = 42) used Blooket for repeated vocabulary retrieval, while the control group (n = 40) received conventional teacher-led review, textbook exercises, oral repetition, and written practice. Pre-test and post-test scores were analysed using paired-samples and independent-samples t-tests, and classroom observations were used to contextualise the learning process. The experimental group improved from M = 5.26 to M = 7.31, $t(41) = 14.59$, $p < .001$, with a very large effect size (Cohen's $d = 2.25$), and outperformed the control group on the post-test. Observation data indicated that repeated retrieval, immediate feedback, peer support, and teacher scaffolding contributed to the gains. The findings support the use of Blooket as a supplementary review tool, provided that game-based practice is followed by speaking and writing tasks that promote productive vocabulary use.

Keywords: AI-powered learning, Blooket, game-based vocabulary practice, vocabulary retention, young EFL learners

1. Introduction

1.1. Background and Context

Vocabulary knowledge is essential to English comprehension and production, yet young EFL learners often have limited opportunities to revisit and use new words beyond the classroom. In many Vietnamese language-centre classes, vocabulary review still relies heavily on teacher explanation, textbook exercises, oral repetition, and short written tasks. These practices remain useful, but they may not provide sufficient active retrieval, immediate feedback, or sustained participation for children aged 10–11. Digital game-based platforms may help address this need by increasing the frequency and intensity of vocabulary practice.

This context matters because private language centres serve a substantial number of Vietnamese children who study English in addition to their mainstream schooling. These learners frequently encounter short lesson periods, mixed-ability classes, and limited English

[1] *Corresponding author, Graduate Student, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology, Ho Chi Minh City, Vietnam; Email: MES24155@ms.huflit.edu.vn*

[2] *Lecturer, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology, Ho Chi Minh City, Vietnam*

exposure outside class. Evidence from this setting can therefore inform teachers who must combine curriculum coverage with motivating, time-efficient vocabulary recycling.

1.2. Problem Statement

Although Blooket is increasingly used to make vocabulary lessons more engaging, classroom enthusiasm does not by itself demonstrate durable learning. Existing studies have often emphasised motivation, enjoyment, or general perceptions, while fewer have compared measurable vocabulary retention outcomes among young Vietnamese EFL learners. It also remains unclear which classroom processes, including repeated retrieval, immediate feedback, peer interaction, and teacher scaffolding, help explain any observed gains. Without such evidence, teachers may adopt game-based tools primarily for their entertainment value without knowing how to integrate them effectively into vocabulary instruction.

1.3. Research Gap

Previous Blooket research has largely focused on university students, older learners, or broad measures of motivation and acceptance (Huynh, 2024; Pham & Ly, 2023; Quyen & Nguyen, 2025; Thu & Dan, 2023). Fewer studies have examined vocabulary retention as the principal outcome among children aged 10–11 in Vietnamese language-centre settings. In addition, limited research has combined pre-test/post-test comparisons with classroom observations to explain how game-based practice operates during instruction. This study addresses these gaps by comparing Blooket-supported and conventional vocabulary review and by using observation data to interpret the quantitative results.

1.4. Research Objectives

The study aimed to:

(1) determine whether Blooket-supported vocabulary practice improves young learners' vocabulary retention;

(2) compare the vocabulary gains of learners receiving Blooket-supported practice with those receiving conventional instruction;

(3) use classroom observations to explain the learning processes associated with the outcomes;

(4) identify practical conditions for effective classroom implementation.

1.5. Research Questions

RQ1. To what extent does Blooket-supported game-based vocabulary practice improve young EFL learners' vocabulary retention compared with conventional vocabulary instruction?

RQ2. How do classroom observations help explain the vocabulary-learning gains observed after the intervention?

RQ3. What pedagogical conditions support the effective integration of Blooket into young EFL vocabulary classrooms?

1.6. Research Hypotheses

H0: There is no statistically significant difference between the vocabulary gains of learners receiving Blooket-supported practice and those receiving conventional vocabulary instruction.

H1: Learners receiving Blooket-supported practice demonstrate significantly greater vocabulary gains than learners receiving conventional vocabulary instruction.

1.7. Scope and Delimitations

The study was delimited to 82 A2-level learners aged 10-11 at one private language centre in Ho Chi Minh City. It examined an eight-week vocabulary intervention using intact classes, with vocabulary retention measured through immediate pre-tests and post-tests. The study focused on Blooket as a supplementary review tool and did not evaluate all language skills, long-term delayed retention, or generative artificial-intelligence tutoring. Classroom observations were limited to the experimental group and were used to contextualise implementation rather than to represent all Vietnamese language-centre classrooms.

1.8. Significance of the Study

The study contributes context-specific evidence on the use of game-based vocabulary practice with young Vietnamese learners. It is practically relevant to teachers and language-centre managers because it identifies not only whether learners improved but also how pacing, feedback, grouping, and follow-up tasks influenced implementation. The findings may support more purposeful decisions about when and how Blooket should be used in vocabulary lessons.

2. Literature Review

2.1. Vocabulary Retention and Repeated Retrieval

Vocabulary retention refers to learners' ability to remember and retrieve lexical items after initial learning. Vocabulary is central to language learning because it supports comprehension and production in listening, speaking, reading, and writing (Nation, 2001). However, knowing a word is not limited to memorizing its meaning; it also includes knowledge of form, use, collocation, and context (Schmitt, 2010). Therefore, vocabulary retention should be understood as the ability to access and use words after repeated learning experiences.

Repeated exposure and retrieval are important for vocabulary development. Learners need to encounter words several times in different contexts to strengthen lexical knowledge (Laufer & Goldstein, 2004; Webb & Nation, 2017). For young EFL learners, repetition should not be purely mechanical because monotonous drills may reduce attention. Instead, vocabulary practice should include interaction, feedback, and meaningful participation, which are suitable for young learners' cognitive and affective characteristics (Cameron, 2001; Pinter, 2006).

2.2. Game-Based Vocabulary Practice

Game-based learning can make vocabulary review more active and engaging. Gamification refers to the use of game elements, such as points, challenges, rewards, and competition, in non-game learning contexts (Deterding et al., 2011; Kapp, 2012). In vocabulary lessons, these elements may encourage learners to review words repeatedly and participate more actively.

However, games do not automatically lead to learning. Their effectiveness depends on instructional design and classroom management (Hamari et al., 2014; Subhash & Cudney, 2018). In vocabulary teaching, game-based activities are useful when they help learners notice word forms, retrieve meanings, receive feedback, and reuse target vocabulary. Therefore, the game should support the learning objective rather than distract students from it.

2.3. Blooket Interface and Vocabulary Practice

Blooket provides a wide range of game modes that integrate learning tasks with playful competition (Huang, 2022). Each mode introduces different challenges and rewards, giving learners multiple opportunities for repeated practice and motivation:

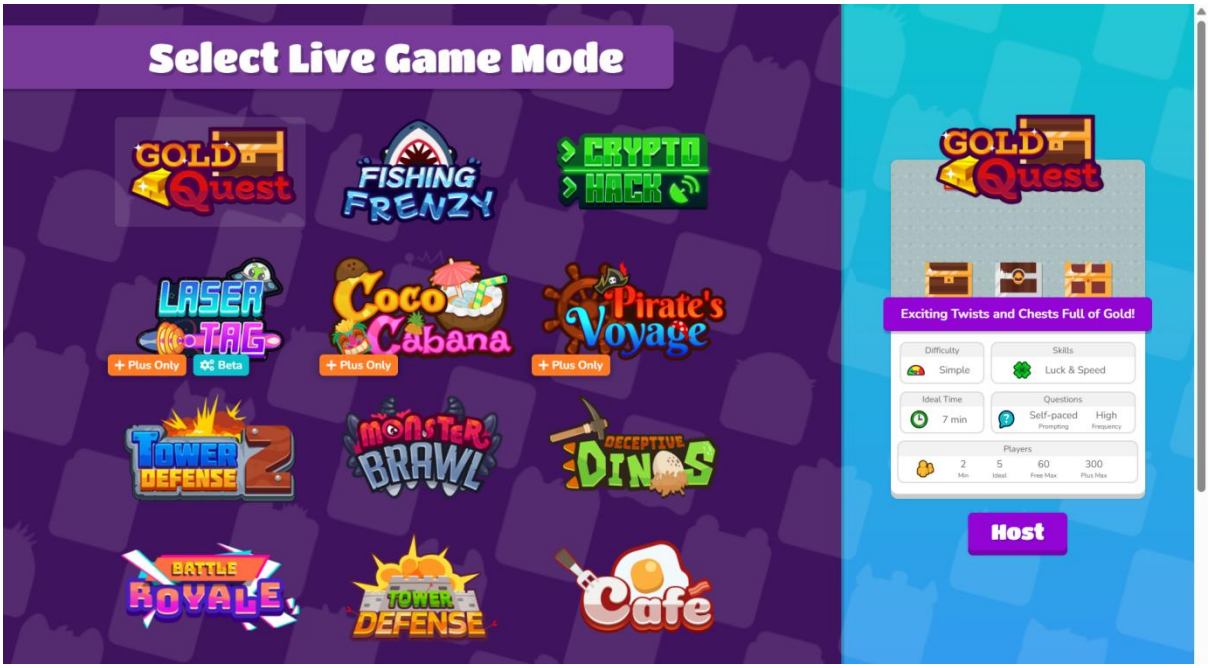


Figure 1

The live game mode options on Blooket

Gold Quest: Players answer questions correctly to open treasure chests, which may contain points or special items, including the ability to steal scores from other players.

Crypto Hack: Learners choose passwords and attempt to "hack" other players' accounts by answering questions and guessing their passwords correctly to gain currency.

Fishing Frenzy: Players catch fish of different weights by responding accurately, with heavier catches yielding higher points.

Blook Rush: Participants can attack or defend. Correct answers allow players to reduce opponents' shields or add protection for their team.

Battle Royale: Learners compete individually or in teams, eliminating others through correct responses until a winner remains.

Tower Defense: Players build and upgrade towers to defend against waves of enemy "blooks," while answering questions to strengthen their defenses.

Café: Students role-play as café staff, answering questions to earn points that are exchanged for serving dishes.

Factory: Players acquire different units that produce resources at varying rates, with correct answers helping them optimize their factories.

Racing: Participants compete to be the fastest to answer all questions correctly, with progress visualized as a race.

Classic: The simplest format where learners answer questions in sequence, similar to a traditional quiz but enhanced with playful animations.

From the teacher's side, Blooket provides automated scoring, instant feedback, response tracking, and class performance data. These features can help teachers identify difficult vocabulary items and plan follow-up instruction. In this study, Blooket is described as AI-powered only in this cautious operational sense, not as a fully adaptive or generative AI system. This view is consistent with research suggesting that automated feedback and learning analytics can support more responsive teaching (Chen et al., 2020; Hwang et al., 2020; Zawacki-Richter et al., 2019).

2.4. Blooket-Supported Vocabulary Retention

Blooket may support vocabulary retention in three main ways. First, it provides repeated encounters with target words through several rounds of gameplay, which is important for vocabulary learning (Nation, 2001; Webb & Nation, 2017). Second, it encourages active retrieval because learners need to choose or recall answers during the game. Third, it gives immediate feedback, allowing learners to notice and correct errors quickly.

Previous studies have reported that Blooket can increase learners' enjoyment, motivation, and classroom participation (Huynh, 2024; Pham & Ly, 2023; Quyen & Nguyen, 2025; Thu & Dan, 2023). Some studies have also found positive effects on vocabulary learning outcomes (Phupunna & Hongsamsibkao, 2024; Sinaga et al., 2025; Sulistyanto & Asyhar, 2024). Nevertheless, enjoyment should not be equated with durable vocabulary learning. Teachers still need to connect Blooket activities with lesson objectives, post-game review, and meaningful vocabulary use.

2.5. Conceptual Framework

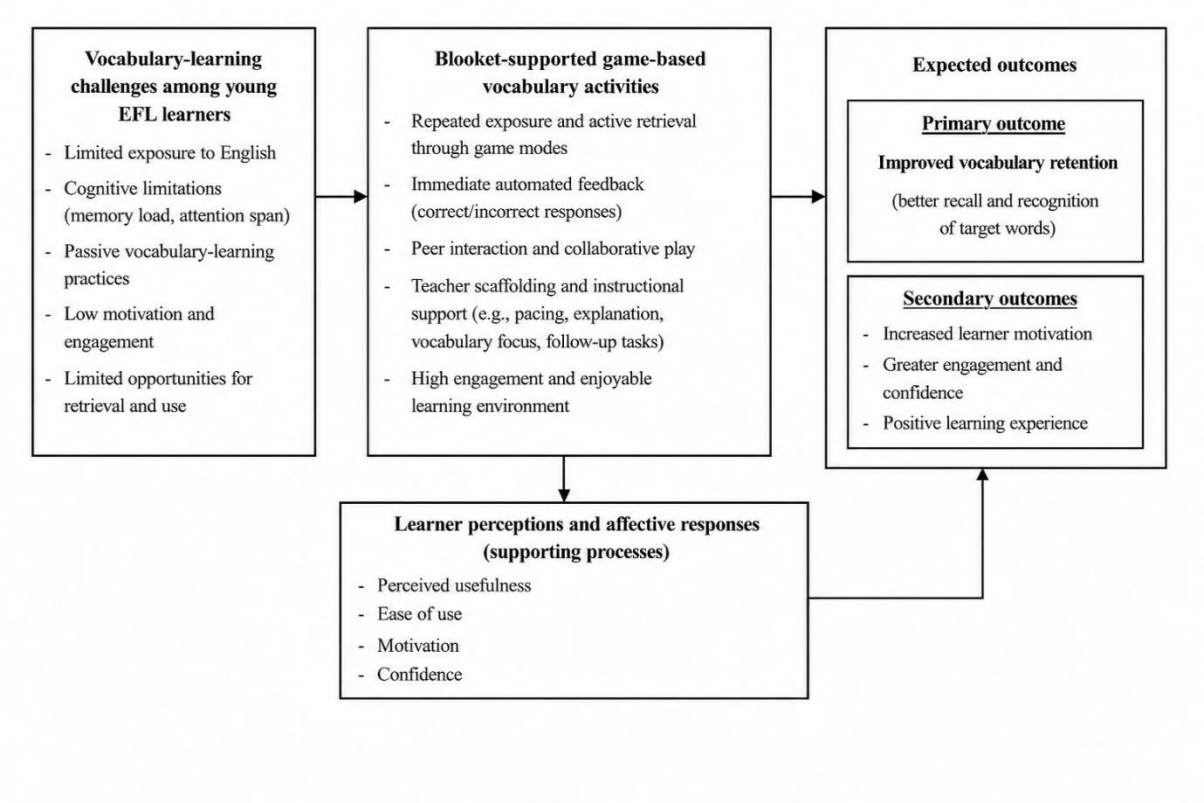


Figure 2

Conceptual Framework of Blooket-Supported Vocabulary Learning among Young EFL Learners

The conceptual framework illustrates how Blooket-supported game-based vocabulary activities may address common vocabulary-learning challenges among young EFL learners. These challenges include limited exposure to English, cognitive limitations, passive vocabulary-learning practices, low motivation, and insufficient opportunities to retrieve and use target words. This context is particularly relevant because young learners require repeated, engaging, and appropriately scaffolded practice to sustain attention and strengthen vocabulary retention.

Blooket-supported activities are proposed to address these challenges through repeated exposure, active retrieval, immediate automated feedback, peer interaction, and teacher scaffolding. Repeated game rounds provide learners with opportunities to encounter and retrieve target vocabulary, while immediate feedback helps them recognise and correct errors. Peer interaction and collaborative play may further support participation, whereas teacher scaffolding, including appropriate pacing, explanation, vocabulary focus, and follow-up tasks, ensures that game participation remains connected to the learning objectives.

The primary expected outcome is improved vocabulary retention, particularly better recall and recognition of target words. Increased motivation, engagement, confidence, and positive learning experiences are treated as secondary outcomes. Learner perceptions and affective responses, including perceived usefulness, ease of use, motivation, and confidence, are included as supporting processes that may help explain learners' participation in Blooket-based activities. However, these factors are not presented as statistically tested mediating variables.

The framework does not assume that Blooket alone produces vocabulary gains. Instead, it proposes that its educational value emerges from the interaction between game-based features, learner responses, peer participation, and teacher guidance. This framework therefore guides the interpretation of both the vocabulary-test results and the classroom-observation evidence in the present study.

3. Methodology

3.1. Research Design

The study used a quasi-experimental classroom design. This design was appropriate because the research was conducted in a real language-centre setting where random assignment of individual learners was not feasible, so intact classes were used (Cook & Campbell, 1979). Quantitative data from vocabulary pre-tests and post-tests were used to examine vocabulary retention outcomes. Classroom observation data were used to contextualize how learners participated during Blooket-supported activities and how teacher scaffolding shaped the implementation.

This paper focuses specifically on vocabulary retention outcomes. Questionnaire and interview data are not foregrounded because the purpose of the paper is to examine measurable vocabulary learning gains and classroom implementation rather than learners' general perceptions of Blooket.

3.2. Context and Participants

The research was conducted at Minds Connect Academy, a private English language centre in Ho Chi Minh City, Vietnam. This setting was selected because private language centres provide substantial supplementary English instruction to Vietnamese children but remain underrepresented in research on game-based vocabulary learning. The participants were 82 A2-level learners aged 10–11, assigned by intact class to an experimental group (n = 42) and a control group (n = 40). The experimental group comprised three classes, coded as SE.9, SE.10, and SE.11 for reporting purposes. Both groups studied the same target vocabulary over the same eight-week period.

The experimental group practiced vocabulary through Blooket-supported game-based activities. The control group received conventional vocabulary instruction, including teacher explanation, textbook-based practice, oral repetition, short written exercises, and classroom review without Blooket. Both groups completed vocabulary pre-tests before the intervention and vocabulary post-tests after the intervention.

3.3. Intervention Procedure

The intervention lasted eight weeks. In the experimental group, Blooket was used as a supplementary vocabulary review tool after the teacher had introduced or practiced the target words. The 42 learners in the experimental condition participated through three separate classes: SE.9, SE.10, and SE.11. Although the classes were organized separately for normal centre scheduling, they followed the same intervention focus, target vocabulary, and general teaching procedure.

Blooket activities were designed to recycle weekly vocabulary through repeated game rounds. Learners were required to recognize meanings, select correct answers, and respond under time constraints. The teacher also used Blooket results and classroom observations to identify vocabulary items that needed further explanation or review. After the game rounds, learners completed short follow-up tasks, such as oral responses, sentence practice, or controlled speaking activities, to extend vocabulary use beyond recognition-level practice. The control group learned the same vocabulary topics through conventional classroom procedures, including teacher-led review, workbook practice, oral repetition, and short written exercises. This comparison made it possible to examine whether the Blooket-supported condition produced stronger vocabulary retention outcomes than regular instruction over the same period.

3.4. Instruments and Data Analysis

The main instruments for this paper were vocabulary pre-tests and post-tests. The tests were adapted from Cambridge-aligned A2 Flyers practice materials and administered under similar classroom conditions. Descriptive statistics were used to compare mean scores and standard deviations. Paired-samples t-tests were used to examine within-group gains from pre-test to post-test. Independent-samples t-tests were used to compare vocabulary performance between the experimental and control groups.

Classroom observations were conducted in the three experimental classes, SE.9, SE.10, and SE.11. Observation notes focused on attention, participation, peer interaction, teacher support, game pacing, and observable vocabulary recall. These qualitative data were not used to make broad claims about perceptions. Instead, they were used to explain how Blooket-supported practice operated during classroom vocabulary review.

3.5. Ethical Considerations

Written permission to conduct the study was obtained from the management of the language centre. Because all participants were minors aged 10–11, parents or legal guardians

were informed about the study through the official parent communication groups before data collection. The notice explained the purpose of the study, the classroom procedures, the voluntary nature of participation, the academic use of the data, and the measures used to protect participants' privacy. Parents or legal guardians were given the opportunity to decline their child's participation.

Learners also received an age-appropriate explanation of the study and provided verbal assent before participating. Participation was voluntary, and learners could decline or withdraw without any effect on their grades, class standing, or access to regular classroom activities. Learners' names were removed from the dataset and replaced with codes, while classroom observation records used the class codes SE.9, SE.10, and SE.11. Electronic data were stored in password-protected folders accessible only to the researcher.

No formal institutional ethics approval number was issued for this centre-based study. In addition, no separate written parental consent form was collected. The study therefore relied on centre authorisation, parental notification with an opt-out opportunity, learner assent, anonymisation, and data-protection procedures. These procedures are reported transparently and are acknowledged as a limitation of the study.

4. Results

4.1. Control Group Vocabulary Results

The descriptive statistics for the control group's vocabulary pre-test and post-test scores are presented in Table 1.

Table 1

Descriptive Statistics of the Control Group's Pre-test and Post-test Scores

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
CG_Vocabulary_Pre_Test	40	6.00	3.00	9.00	5.575	1.412	1.994
CG_Vocabulary_Post_Test	40	6.00	3.00	9.00	6.075	1.509	2.276

As shown in Table 1, the control group showed a small increase in vocabulary performance after the eight-week learning period. The mean score rose from 5.575 (SD = 1.412) in the pre-test to 6.075 (SD = 1.509) in the post-test. The minimum and maximum scores remained unchanged at 3.00 and 9.00 respectively. These results suggest that learners who received conventional vocabulary instruction made some progress, but the improvement was modest.

Table 2

Test of Normality for the Control Group's Vocabulary Pre-test and Post-test Scores

	Shapiro-Wilk		
	Statistic	df	Sig.
CG_Vocabulary_Pre_Test	.953	40	**.097**
CG_Vocabulary_Post_Test	.957	40	**.136**

As presented in Table 2, the Shapiro-Wilk significance values for the control group were greater than .05 for both the pre-test (p = .097) and the post-test (p = .136). Therefore, the control group's vocabulary scores met the assumption of normality.

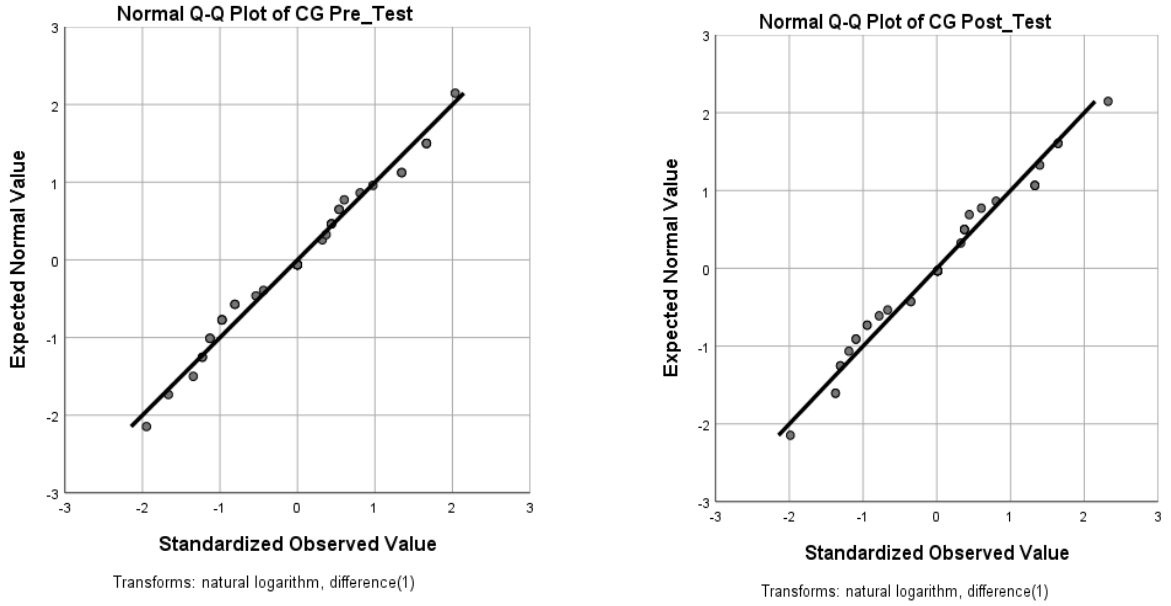


Figure 3

Normal Q-Q Plots for the Control Group's Vocabulary Test Results

The Normal Q-Q plots in Figure 3 also support this conclusion because the data points generally follow the diagonal line. Taken together, the statistical and graphical checks indicate that the control group's pre-test and post-test scores were suitable for parametric analysis.

4.2. Experimental Group Vocabulary Results

The descriptive statistics for the experimental group's vocabulary pre-test and post-test scores are presented in Table 3.

Table 3

Descriptive Statistics of the Experimental Group's Pre-test and Post-test Scores

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
EG_Vocabulary_Pre_Test	42	9.00	1.00	10.00	5.261	2.084	4.344
EG_Vocabulary_Post_Test	42	7.00	3.00	10.00	7.309	1.675	2.804

As shown in Table 3, the experimental group showed a clear increase in vocabulary performance after the Blooket-supported intervention. The mean score increased from 5.261 (SD = 2.084) in the pre-test to 7.309 (SD = 1.675) in the post-test. The minimum score also increased from 1.00 to 3.00, while the maximum score remained 10.00. The higher mean score and lower standard deviation in the post-test suggest that learners not only improved overall but also became more consistent in their vocabulary performance.

Table 4

Test of Normality for the Experimental Group's Vocabulary Pre-test and Post-test Scores

	Shapiro-Wilk		
	Statistic	**df**	**Sig.**
EG_Vocabulary_Pre_Test	.972	42	.378
EG_Vocabulary_Post_Test	.954	42	.090

As presented in Table 4, the Shapiro-Wilk significance values were greater than .05 for both the pre-test (p = .378) and the post-test (p = .090). This indicates that the experimental group's vocabulary scores were normally distributed and met the assumption required for a paired-samples t-test.

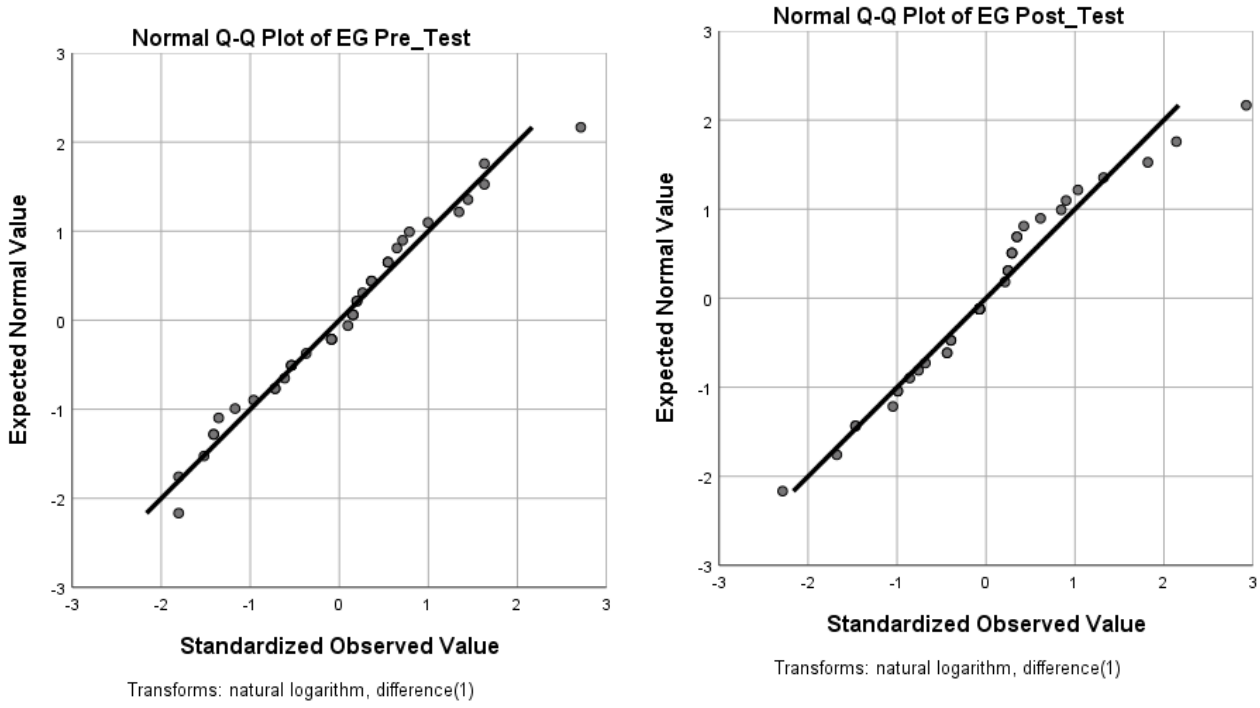


Figure 4

Normal Q-Q Plots for the Experimental Group's Vocabulary Test Results

The Q-Q plots in Figure 4 further confirm the normality assumption because the observed values are generally aligned with the diagonal reference line. Therefore, both the pre-test and post-test scores of the experimental group were appropriate for parametric comparison.

4.3. Between-Group Comparison of Pre-test Scores

Before the intervention, both groups completed the vocabulary pre-test to determine whether their initial vocabulary knowledge was comparable. The control group had a mean score of 5.58 (SD = 1.41), while the experimental group had a mean score of 5.26 (SD = 2.08).

Table 5

Descriptive Statistics for the Control and Experimental Groups' Pre-test Scores

Group	N	Range	Minimum	Maximum	M	SD
Control group	40	6.00	3.00	9.00	5.58	1.41

Group	N	Range	Minimum	Maximum	M	SD
Experimental group	42	9.00	1.00	10.00	5.26	2.08

Because Levene's test was significant, F = 5.827, p = .018, the unequal-variances result was interpreted. The independent-samples t-test indicated no statistically significant difference between the pre-test scores of the control and experimental groups, t(72.37) = 0.80, p = .427. Therefore, the two groups began the intervention with comparable vocabulary knowledge.

Table 6

Independent-Samples t-test for Vocabulary Pre-test Scores

Variance assumption	Levene F, p	t (df)	p	Mean difference (SE)	95% CI
Equal variances assumed	5.827, .018	.792 (80)	.554	.313 (.395)	[-.473, 1.099]
Equal variances not assumed	—	.800 (72.373)	.427	.313 (.392)	[-.467, 1.094]

4.4. Between-Group Comparison of Post-test Scores

After the intervention, both groups completed a vocabulary post-test with the same format and level of difficulty as the pre-test. The experimental group obtained a higher mean score (M = 7.31, SD = 1.67) than the control group (M = 6.08, SD = 1.51). An independent-samples t-test was conducted to determine whether this difference was statistically significant.

Table 7

Independent-Samples t-test for Vocabulary Post-test Scores

Variance assumption	Levene F, p	t (df)	p	Mean difference (SE)	95% CI
Equal variances assumed	.507, .479	-3.501 (80)	.001	-1.235 (.353)	[-1.936, -.533]
Equal variances not assumed	—	-3.510 (79.761)	.001	-1.235 (.352)	[-1.934, -.535]

Levene's test was not significant, F = 0.507, p = .479, indicating that the equal-variances assumption was met. The independent-samples t-test showed a statistically significant difference between the post-test scores of the two groups, t(80) = -3.50, p = .001. The experimental group scored significantly higher than the control group, with a mean difference of 1.23 points. This result supports the conclusion that the Blooket-supported condition produced stronger vocabulary outcomes than conventional review during the eight-week intervention.

4.5. Paired-Samples T-test Results for the Experimental Group

After confirming the normality of the experimental group's test scores, a paired-samples t-test was conducted to determine whether the increase from pre-test to post-test was statistically significant. This analysis directly addressed whether Blooket-supported game-based vocabulary practice contributed to measurable vocabulary gains after the eight-week intervention.

Table 8

Results of the Paired-Samples T-test for the Experimental Group

	Mean	N	Std. Deviation	Std. Error Mean
Pre-test	5.26	42	2.08	.32
Post-test	7.31	42	1.67	.26

	Mean Difference	Std Error Mean	95% CI		t	df	Sig. (2-tailed)	Effect Size (Cohen's *d*)
			Lower	Upper				
Post-test - Pre-test	2.05	.140	1.69	2.27	14.59	41	< .001	2.25

The paired-samples t-test showed that the experimental group made a statistically significant improvement in vocabulary performance. As shown in Table 8, the mean score increased from 5.26 (SD = 2.08) in the pre-test to 7.31 (SD = 1.67) in the post-test. The mean gain was 2.05 points, and the test result was statistically significant, $t(41) = 14.59$, $p < .001$. This indicates that students performed significantly better after the Blooket-supported vocabulary practice than before the intervention.

The effect size was very large (Cohen's d = 2.25), suggesting that the observed gain was not only statistically significant but also educationally meaningful. However, this result should be interpreted as evidence that Blooket-supported practice was effective as part of the classroom vocabulary review cycle, rather than as evidence that the platform alone caused all improvement.

According to Cohen's (1988) conventional interpretation, effect sizes of approximately 0.20, 0.50, and 0.80 may be considered small, medium, and large respectively. The value obtained in this study was well above the large-effect threshold, which supports the practical value of integrating Blooket into vocabulary review when it is guided by teacher explanation, feedback, and follow-up practice.

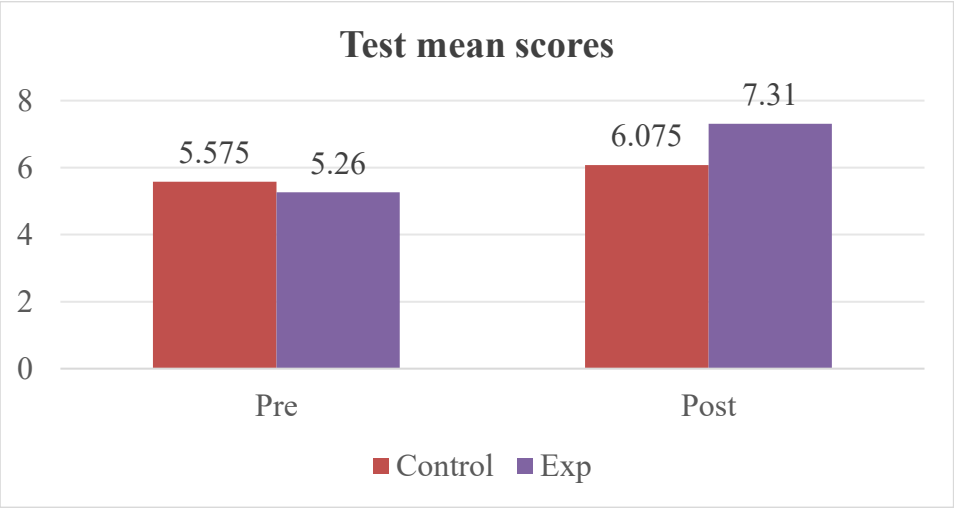


Figure 5

Comparison of Mean Pre-test and Post-test Scores

As illustrated in Figure 5, both groups improved from pre-test to post-test, but the increase was much larger in the experimental group. The control group increased by 0.50 points, whereas the experimental group increased by approximately 2.05 points. This pattern is consistent with the paired-samples t-test result and suggests that Blooket-supported vocabulary practice provided stronger support for vocabulary retention than conventional review alone.

4.6. Classroom Observation

Classroom observations were used to contextualize the quantitative test results and to explain how the Blooket-supported vocabulary practice operated during the intervention. The observations focused on students' attention, participation, peer interaction, vocabulary recall, teacher support, and classroom atmosphere in the experimental group at Minds Connect Academy.

The experimental group consisted of 42 learners divided into three centre-based classes, coded as SE.9, SE.10, and SE.11 for observation purposes. Each class received the same Blooket-supported vocabulary review procedure during the eight-week intervention. The observation checklist was created to capture critical elements of implementation, such as teacher facilitation, student engagement, interactions, pace, and vocabulary recall. During the observed lessons, Blooket was used mainly to review and recycle vocabulary previously introduced by the teacher.

In Class SE.9, students showed high engagement during Blooket vocabulary activities. After a brief teacher-led review, learners completed repeated game rounds and frequently discussed answers with peers. Several students participated more readily than during conventional review, and many were able to reuse target words in short, controlled speaking tasks after the game.

In Class SE.10, scores and leaderboards helped most learners maintain attention. However, some lower-proficiency learners struggled with faster-paced game modes. The teacher responded by pausing the activity, clarifying difficult words, adjusting the pace, and encouraging peer support. This observation suggests that Blooket was most effective when the teacher actively regulated pacing and provided timely scaffolding.

In Class SE.11, learners were placed in mixed-ability groups to reduce pressure and encourage cooperation. Students supported one another with pronunciation, word meanings, and answer selection. This grouping appeared to broaden participation and supported vocabulary recall during the speaking activities that followed the game.

Overall, the observations were consistent with the quantitative improvement. Repeated retrieval, immediate feedback, peer interaction, and sustained attention appeared to support vocabulary learning. At the same time, the observations showed that Blooket worked best as guided supplementary practice rather than as a substitute for teacher explanation or follow-up speaking and writing tasks.

5. Discussion

The findings indicate that Blooket-supported vocabulary practice was associated with stronger learning gains than conventional review in this classroom context. The two groups did not differ significantly at pre-test, $t(72.37) = 0.80$, $p = .427$, but the experimental group achieved a significantly higher post-test score than the control group, $t(80) = -3.50$, $p = .001$. Within the experimental group, the mean increased from $M = 5.26$ to $M = 7.31$, $t(41) = 14.59$, $p < .001$, with a very large effect size (Cohen's $d = 2.25$). Because the study used intact classes

rather than random assignment, these findings should be interpreted as strong classroom-based evidence rather than a universal causal claim.

The observed gains are consistent with research emphasising repeated encounters and active retrieval in vocabulary learning (Nation, 2001; Webb & Nation, 2017). Blooket required learners to retrieve target meanings several times within a short period, while immediate feedback allowed them to notice and correct errors during practice. These features likely made vocabulary recycling more intensive than in conventional review alone.

Classroom observations clarified how the intervention operated. Learners remained attentive, discussed answers, supported peers, and responded to teacher feedback. The benefit therefore appeared to arise from the interaction of technology and pedagogy rather than from the game format alone. Teacher explanation before play, monitoring during activities, and follow-up language use were central to the instructional sequence.

The observations also showed that game modes were not equally accessible to all learners. Faster-paced competition increased excitement for some students but placed additional pressure on lower-proficiency learners. Teachers should therefore adjust timing, pause for clarification, use mixed-ability grouping, and select modes according to learners' needs rather than assuming that speed and competition are universally motivating.

Blooket was particularly effective for recognition and recall, but these outcomes do not automatically represent productive vocabulary knowledge. Learners still require opportunities to pronounce words, build sentences, answer questions, and use vocabulary in speaking or writing. Game-based review should therefore be followed by tasks that move learners from rapid answer selection to meaningful language production.

The term AI-powered should also be interpreted cautiously. In this study, it refers to automated scoring, immediate feedback, response tracking, and teacher-accessible performance data, not to generative or fully adaptive artificial intelligence. These functions supported instructional decisions, but they did not replace teacher judgement, lesson design, or classroom management.

The study adds context-specific evidence from a Vietnamese private language centre, a setting in which learners often have limited exposure to English outside class and teachers must use lesson time efficiently. The findings suggest that Blooket can enrich conventional vocabulary instruction when it is integrated into a structured cycle of explanation, retrieval, feedback, collaboration, and productive follow-up practice.

6. Pedagogical Recommendations

The pedagogical recommendations are directly based on the findings of the study. Because the test and observation findings suggested stronger support for vocabulary review and recall than for productive sentence use, teachers should introduce or review target words before the activity and follow each game with short speaking or writing tasks. These tasks may include sentence construction, picture description, dialogue completion, or pair questioning so that learners can move from recognition and recall to productive vocabulary use.

Because the classroom observations showed that fast-paced modes were difficult for some lower-proficiency learners, teachers should select game modes according to learners' proficiency levels and processing speed. Slower or team-based formats may be more appropriate for mixed-ability classes. Mixed-ability grouping and peer support can also reduce pressure and encourage wider participation.

During gameplay, teachers should monitor common errors, pause when learners experience difficulty, and provide further explanation of challenging vocabulary. Blooket response data can also be used to identify items that require additional review. As the findings showed that the effectiveness of Blooket depended on teacher guidance, the platform should be used as a supplementary tool rather than as a replacement for explicit vocabulary instruction.

Language-centre managers should provide teachers with practical guidance on selecting suitable game modes, controlling activity pacing, protecting learner data, and aligning Blooket activities with curriculum objectives. Such support can help ensure that game-based activities remain inclusive, purposeful, and connected to meaningful language learning.

7. Limitations and Future Research

The study was conducted with intact classes at one private language centre, which limits generalisability and does not provide the control associated with random assignment. The intervention lasted eight weeks and used an immediate post-test, so the results do not establish long-term retention. Classroom observations were conducted only in the experimental group and were interpreted by the teacher-researcher, creating possible observer and expectancy bias. The study also focused mainly on recognition and recall rather than broader productive vocabulary knowledge. Future research should include multiple sites, delayed post-tests, independent observers, and tasks measuring sentence-level and communicative vocabulary use.

The study did not use a separate written parental consent form. Instead, parents or legal guardians were informed through the official parent communication groups and were given the opportunity to decline their child's participation. Although this procedure was consistent with the centre's usual communication practices, the absence of documented individual parental consent limits the strength of the study's ethical documentation. Future studies involving minors should obtain formal written informed consent from parents or legal guardians before data collection.

8. Conclusion

This study examined the effects of Blooket-supported game-based practice on vocabulary retention among young Vietnamese EFL learners. The experimental group achieved substantially greater gains than the control group, and observation data linked these gains to repeated retrieval, immediate feedback, peer support, and teacher scaffolding.

Blooket should nevertheless be treated as a supplementary instructional tool rather than a replacement for systematic vocabulary teaching. Its value depends on appropriate pacing, inclusive grouping, teacher mediation, and follow-up speaking and writing activities. Used in this way, Blooket can provide an engaging and instructionally purposeful form of vocabulary recycling for young EFL learners.

Acknowledgments

The author would like to express sincere gratitude to Dr. Nguyen Ho Phuong Chi and Assoc. Prof. Dr. Nguyen Ngoc Vu for their valuable guidance, academic support, and encouragement throughout the research process. This paper was developed as part of a larger research thesis conducted during the author's Master's program in English Language Studies at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). The author also wishes to thank the students at Minds Connect Academy for their participation and support during the classroom implementation.

References

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning future media environments* (pp. 9–15). Association for Computing Machinery. https://doi.org/10.1145/2181037.2181040

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Huang, W. (2022). Blooket: Game on, students! *FLTMAG*. https://fltmag.com/blooket-game-on/

Huynh, D. P. L. (2024). The effects of Blooket on motivation in learning English among first-year non-English majors at a university in Ho Chi Minh City. *International Journal of English Language Education, 12*(1). https://doi.org/10.5296/ijele.v12i1.21758

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Pfeiffer.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399–436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Pham, A. T., & Ly, D. T. (2023). EFL students' perceptions on the use of Blooket in grammar classes. In *Proceedings of the 2023 8th International Conference on Distance Education and Learning*. Association for Computing Machinery. https://doi.org/10.1145/3606094.3606117

Phupunna, S., & Hongsamsibkao, P. (2024). Enhancing Chinese vocabulary memorization skills through Blooket games combined with active learning for first-year students in the Chinese language teaching program at Rajabhat Mahasarakham University. *Journal of Education and Learning, 13*(5), 191–198. https://doi.org/10.5539/jel.v13n5p191

Pinter, A. (2006). *Teaching young language learners*. Oxford University Press.

Quyen, C. T., & Nguyen, H. C. (2025). Perceptions and dominant factors of Blooket as a game-based learning tool among Vietnamese non-English majored students. *Dinamika Ilmu, 25*(1), 123–144. https://doi.org/10.21093/di.v25i1.10268

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan. https://doi.org/10.1057/9780230293977

Sinaga, N., Widyantoro, A., & Sinaga, A. N. (2025). The effectiveness of Blooket on students' vocabulary learning among secondary school students. *Indonesian Research Journal in Education, 9*(2), 1273–1288. https://online-journal.unja.ac.id/irje/article/view/44887

Subhash, S., & Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior, 87*, 192–206. https://doi.org/10.1016/j.chb.2018.05.028

Sulistyanto, I., & Asyhar, W. I. (2024). The effectiveness of Blooket media toward students' vocabulary mastery. *English Education: English Journal for Teaching and Learning, 12*(2), 222–237.

Thu, T. T. M., & Dan, N. T. T. (2023). Students' perceptions on English vocabulary teaching and learning by using Blooket: A case study. *European Journal of Applied Linguistics Studies, 6*(1). https://doi.org/10.46827/ejals.v6i1.321

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Bionote

Trang Trinh Khanh Lien is a graduate student in English Language Studies at Ho Chi Minh City University of Foreign Languages and Information Technology, Vietnam. Her research interests include technology-assisted language learning, young learner vocabulary development, game-based learning, and technology-enhanced classroom practice.

Nguyen Ho Phuong Chi is a lecturer at Ho Chi Minh City University of Foreign Languages and Information Technology and holds a PhD in TESOL from RMIT University, Australia. She has more than 15 years of experience in English language teaching and teacher education, having lectured at both Ho Chi Minh City University of Education and Ho Chi Minh City University of Foreign Languages and Information Technology. Her teaching and research interests include TESOL, language teacher education, curriculum development, and technology-enhanced language learning. She is committed to supporting English language learners and contributing to the advancement of English language teaching and learning in Vietnam through research, professional development, and innovative pedagogical practices.

#100 CREATING ENGLISH-RICH LEARNING ENVIRONMENT BEYOND THE CLASSROOM

Pham Thi Bich Tien[1]

Abstract

This paper addresses a persistent gap in Vietnamese university students' English-speaking development: the disconnect between classroom-based instruction and real-world communicative competence. While prior TESOL research has examined out-of-class learning broadly, this study contributes a distinctively integrated, multi-level intervention framework tailored to the Vietnamese university context, combining psychological safety design, task-based output routines, and institutional alignment mechanisms. Through thematic analysis of recurring barriers documented in the SLA literature, three interlocking themes are identified: psychological inhibition, methodological passivity, and structural disconnection. Based on these themes, the paper proposes a coordinated system of solutions across four levels — individual, peer group, institutional, and community — and presents an 8-week implementation roadmap grounded in evidence-based principles of output-driven language learning. The paper concludes that an English-rich environment beyond classroom activities can effectively support learners in transforming input knowledge into communicative output, thereby building confidence and sustainable speaking development.

Keywords: English-rich environment, beyond classroom, 8-week implementation roadmap, speaking development

1. Introduction

1.1 Background of Study

The era of globalization and rapid digital advancement has positioned English as an indispensable skill for university students. Beyond its role in accessing academic materials, scholarly resources, and research publications, English proficiency has a direct bearing on students' professional futures, shaping their performance in internships, job interviews, and daily workplace communication. Furthermore, English functions as a gateway through which students can engage with academic conferences, seminars, and international exchange programs, ultimately fostering their ability to thrive in diverse and multicultural settings. Given these realities, developing English competence should not be confined to fulfilling classroom obligations or meeting course requirements. Rather, it must be understood as a lifelong, ongoing journey that continues well beyond the boundaries of formal instruction.

Yet in practice, English learning in many universities remains largely confined to the classroom. This gives rise to a significant imbalance between the limited hours of formal language instruction and the far longer stretches of daily life in which English goes largely unused. Without sufficient opportunities to apply the language in authentic, real-world contexts, students easily fall into what is commonly referred to as “learning without using.” Many can follow English videos or comprehend written texts to a reasonable degree yet find

[1] *Faculty of Business Administration, Email: tienptb@huflit.edu.vn*

themselves at a loss when required to speak on the spot, respond swiftly in live conversations, or articulate their thoughts clearly under the natural pressure of real communication. Compounding this challenge is the psychological barrier that many learners face, the fear of making errors and the anxiety of being evaluated by others, which can silence students even when they genuinely have something to say. Consequently, their spoken communication skills progress slowly, not necessarily due to a complete absence of exposure, but rather because meaningful, consistent practice of English as a living, active skill remains critically lacking.

It is for this reason that adopting an approach centered on building an English-rich learning environment beyond the classroom is both highly reasonable and practically sound. Language acquisition, by its very nature, depends on repeated exposure and meaningful use across a wide variety of contexts. When English is encountered only within classroom hours, learners are unlikely to receive sufficient input to develop authenticity and informality in the language, or enough authentic interaction to cultivate the confidence and fluency that real communication demands. On the other hand, when students take deliberate steps to create a rich English learning environment outside the classroom, English gradually becomes woven into the fabric of their everyday lives.

1.2. Problem Statement

Despite the recognized importance of English proficiency in academic and professional contexts, a substantial gap persists between classroom-based language instruction and the communicative competence that students are expected to demonstrate in real-world settings. Many university students complete years of formal English education yet continue to struggle with spontaneous speaking, interactive conversation, and confident oral expression. This gap is not simply a product of inadequate teaching within the classroom; rather, it reflects a systemic absence of structured, meaningful English use outside of formal instruction. The lack of an English-rich extracurricular environment means that students have insufficient opportunities to reinforce, apply, and internalize the language skills they acquire in class. Without deliberate efforts to extend English learning beyond the classroom, students' speaking proficiency and willingness to communicate are unlikely to develop to the level required for academic success and professional readiness.

1.3. Research Objectives

This study aims to achieve the following objectives:

1. To identify the key barriers that prevent university students from practicing English outside the classroom, with particular focus on speaking proficiency and willingness to communicate.
2. To examine the theoretical foundations of extracurricular language learning environments and their role in second language acquisition.
3. To propose a structured, multi-level system of practical strategies for building an English-rich learning environment beyond the classroom.
4. To design a feasible 8-week implementation roadmap that enables students to progressively develop their speaking skills and communicative confidence through extracurricular engagement.

1.4. Research Questions

This study is guided by the following research questions:

1. What are the primary obstacles that hinder university students from using English outside the classroom, particularly in terms of speaking and communication?
2. How can theoretical principles of second language acquisition inform the design of an extracurricular English-rich learning environment?
3. What multi-level strategies, encompassing individual habits, peer interaction, institutional support, and community engagement, can effectively enhance students' English speaking proficiency and willingness to communicate beyond the classroom?
4. How can these strategies be structured into a coherent and practical implementation plan suitable for university students?

1.3. Hypotheses

Based on the research objectives and questions above, this study proposes the following hypotheses:

- H1: University students who are exposed to a deliberately constructed English-rich environment outside the classroom will demonstrate greater improvement in speaking proficiency compared to those who rely solely on classroom instruction.
- H2: Students who engage in structured extracurricular English activities, such as English-speaking clubs, media consumption in English, and peer conversation practice, will report higher levels of willingness to communicate in English.
- H3: A multi-level approach to building an extracurricular English environment, incorporating individual, peer, institutional, and community dimensions, will be perceived by students as more feasible and motivating than single-level interventions.

1.4. Scope and Delimitation of the Study

This study focuses specifically on non-English-major university students whose English learning is primarily delivered through general English courses as part of their academic program. The central concern of this paper is the development of speaking proficiency and willingness to communicate (WTC) in English, rather than the full spectrum of language skills (reading, writing, listening). While reading, listening, and writing activities may be discussed as supporting components of an English-rich environment, they are treated as subsidiary to the primary focus on oral communication.

The proposed strategies and implementation roadmap are designed for application within a Vietnamese university context, drawing on the specific sociolinguistic and institutional realities that characterize English learning in Vietnam. Accordingly, the findings and recommendations of this study may not be directly transferable to other educational or cultural contexts without appropriate adaptation.

This study does not constitute an empirical investigation. No primary data was collected through surveys, interviews, or experiments. The paper is theoretical and propositional in nature, drawing on established second language acquisition theory and relevant literature to construct a reasoned framework of recommendations. Consequently, claims regarding the

effectiveness of the proposed strategies are inferential rather than empirically validated, and future empirical studies are recommended to evaluate the hypotheses presented herein.

1.5. Organization of the Study

This paper is organized into the following chapters: Chapter 1 (Introduction) introduces the background, problem statement, research objectives, research questions, hypotheses, and scope of the study. Chapter 2 (Theoretical Foundation) presents a review of relevant literature on second language acquisition and extracurricular language learning. Chapter 3 (Thematic Analysis of Barriers in the University Context) proposes the multi-level system of strategies for building an English-rich environment beyond the classroom. Chapter 4 (Solutions) outlines the 8-week implementation roadmap. Chapter 5 (Conclusion, Recommendations, and Implementation Roadmap) provides the conclusion and recommendations for future research.

2. Theoretical Foundation

The Concept of a Language Learning Environment

In research on language teaching and learning, the notion of a "language learning environment" is not limited to physical classroom space. Instead, it is understood as an overall system of interacting factors that influence how learners acquire and use a language. A language learning environment includes five key components: resources, interacting with people, learning activities, language-use norms, and motivation (Dörnyei, 2005; MacIntyre & Vincze, 2017).

First, resources refer to the sources of input that learners can access, such as reading materials, videos, podcasts, textbooks, online posts, and academic documents. When resources are diverse and appropriate to learners' proficiency levels, students are more likely to sustain interest and encounter the target language across multiple contexts (VanPatten, 1996; Ellis, 2003).

Second, interacting people are individuals who provide real "dialogue partners" for learners. These can include classmates, instructors, event organizers, members of an online community, or speakers encountered in digital spaces. When learners interact with others, language becomes functional and meaningful; it is used to achieve communication goals rather than to complete exercises (Swain, 1985, 1995).

Third, learning activities are the tasks and interactional practices in which learners take part. An environment is truly "rich" only when learners do not rely solely on passive consumption (e.g., watching or reading) but also engage in purposeful activities such as discussion, responding to others, presenting, asking questions, writing, and producing language-based outputs (Ellis, 2003; Swain, 1985).

Fourth, language-use norms refer to how language is used in real situations, including politeness levels, contextual phrasing, appropriate word choices, and language patterns depending on relationships between speakers. When these norms are reflected in learning situations, learners gradually develop more accurate and natural communicative competence (Swain, 1985; Ellis, 2003).

Finally, motivation determines whether learning efforts can be sustained. Motivation may come from academic needs, career goals, or the sense of personal progress. Effective language learning environments often support motivation by enabling learners to experience success in incremental steps, which is closely related to learners' L2 confidence and willingness to communicate (MacIntyre et al., 1998; Horwitz et al., 1986).

The Role of Learning Beyond the Classroom

Learning English outside the classroom plays a crucial role because it provides opportunities that classroom time often cannot cover adequately. First, out-of-class learning increases the frequency of exposure to real language. While classroom instruction is limited by schedules, students' daily lives and online environments offer continuous streams of language through social media, academic posts, videos, news, and everyday interactions (VanPatten, 1996; Dörnyei, 2005).

Second, learning beyond class enables students to develop skills in a situational and integrated manner. Instead of practicing skills in isolation, learners can connect listening, reading, speaking, and writing to meaningful contexts. For instance, listening to discipline-related podcasts may support note-taking and later writing tasks; participating in online group discussions may improve learners' ability to respond and negotiate meaning; reading short research summaries may provide language input for writing paraphrases or abstracts (Ellis, 2003; Swain, 1985).

Third, learning outside helps reduce fear of making mistakes and increases confidence. When learners are offered a psychologically safe space to experiment with language, sometimes in informal settings such as clubs, discussion groups, or learning journals, they experience less pressure to be judged for errors. As a result, learners begin to use English as a communication tool rather than only as an object for testing (Horwitz et al., 1986; MacIntyre & Vincze, 2017).

Principles for Creating an Effective Language Learning Environment

Based on foundational ideas in modern language education, an effective language learning environment can be built on the following principles:

- Comprehensible input: Learners should receive language input at a level that is challenging but manageable (VanPatten, 1996). Krashen's (1982) Input Hypothesis similarly argues that acquisition occurs when input is slightly beyond the learner's current competence. Selecting contextually appropriate content is therefore essential to sustaining both comprehension and motivation.
- Meaningful interaction: Learners need opportunities to interact so that they can clarify meaning, negotiate communication, confirm information, and respond. Long's (1983) Interaction Hypothesis demonstrates that modified interaction, where communication breakdowns are negotiated, provides learners with comprehensible input and feedback essential for development. Without interaction, learners may remain confined to passive comprehension (Swain, 1985; Ellis, 2003).
- Pushed output: A rich environment does not only provide input; it must also promote output. Swain's (1985, 1995) Output Hypothesis argues that when learners are required to produce language to express precise meaning, they notice gaps in their interlanguage, test hypotheses about the target language, and develop greater syntactic control. Output thus strengthens both accuracy and communicative confidence.
- Continuity and gradual difficulty increase: Exposure and practice should be maintained regularly to avoid learning only during peak periods. As learners become comfortable with a certain level of tasks, difficulty should gradually

increase, for example, by extending task length, increasing processing speed, or raising task openness (Skehan & Foster, 1999; Ellis, 2009).

- Task-based learning: Learners make better progress when they engage in tasks with clear goals and tangible outcomes. Ellis (2003, 2009) identifies task-based language teaching as particularly effective because it creates conditions where language use is purposeful, meaningful, and authentically communicative, precisely the conditions absent from traditional classroom-bound instruction.
- Purposeful learning: Learning objectives should be individualized and connected to concrete needs. When learners understand why they are learning, intrinsic motivation improves and effort becomes more consistent (Dörnyei, 2005; MacIntyre et al., 1998).

From these principles, it can be concluded that creating an English-rich environment beyond the classroom is not merely an extracurricular or symbolic activity. It should be treated as an intentional educational design that includes appropriate input, meaningful interaction, required output, task-based goals, and continuous follow-up.

Thematic Analysis of Barriers in the University Context

A thematic analysis of the barriers facing Vietnamese university students in developing English speaking proficiency reveals three interrelated themes: psychological inhibition, methodological passivity, and structural disconnection. These themes do not operate in isolation; rather, they form a mutually reinforcing cycle that this paper terms the non-use loop. Each theme is discussed below with reference to empirical and theoretical evidence.

Theme 1: Psychological Inhibition – Fear, Anxiety, and Performance Pressure

The most immediate and visible barrier is the cluster of psychological factors that suppress learners' willingness to communicate (WTC). Horwitz and Cope (1986) identified foreign language classroom anxiety (FLCA) as a distinct, complex psychological construct tied specifically to language learning contexts. Their research demonstrated that FLCA comprises three components: communication apprehension, test anxiety, and fear of negative evaluation, all of which are routinely activated in speaking situations.

In the Vietnamese university context, this anxiety is compounded by sociolinguistic factors. Many students have been educated in a predominantly grammar-translation tradition, in which accuracy is prioritized over fluency and errors carry social consequences (Nguyen, 2011). Consequently, speaking in English becomes associated not with communication but with evaluation and potential embarrassment. MacIntyre et al.'s (1998) situational model of WTC demonstrates that when learners anticipate negative outcomes, such as social judgment or loss of face, their communicative intention decreases, even when linguistic knowledge is sufficient.

A further dimension is the performance-oriented mindset: learners focus on producing error-free language rather than prioritizing communicative effectiveness. This leads to over-monitoring, cognitive overload, and disfluency under time pressure. The result is not a language deficit but a confidence deficit; learners know but cannot deploy what they know under real communicative conditions.

Theme 2: Methodological Passivity – Input without Output

The second theme concerns the predominance of receptive, input-focused learning strategies among university students. Common patterns include watching English videos,

reading online content, and memorizing vocabulary, all of which are valuable but insufficient if not accompanied by productive output.

Swain's (1985, 1995) Output Hypothesis provides the theoretical foundation for understanding why input-heavy methods are insufficient. Comprehensible output, language produced by the learner to convey meaning, serves three key functions: it forces learners to notice gaps in their interlanguage; it provides opportunities to test linguistic hypotheses; and it promotes automaticity through repeated production practice. None of these functions is activated through passive reception alone.

Empirical evidence supports this view. Izumi (2002) found that learners who produced output after reading showed significantly greater noticing of target forms than those who read alone. Similarly, Kowal and Swain (1994) demonstrated that collaborative output tasks led learners to reflect on their language use in ways that input tasks did not. These findings suggest that students who rely predominantly on input activities are systematically depriving themselves of the cognitive processes most essential for speaking development.

A further methodological issue is the absence of long-term, cyclical planning. Learning tends to be concentrated around exam periods, with extended intervals of minimal practice, a pattern inconsistent with what is known about distributed practice and memory consolidation (Dörnyei, 2005). Without a structured schedule, even motivated learners struggle to convert knowledge into durable communicative competence.

Theme 3: Structural Disconnection – The Gap between Classroom and Real-World Use

The third theme addresses a systemic problem: the structural absence of pathways connecting classroom learning to authentic out-of-class practice. Many universities treat extracurricular English activities as optional add-ons rather than integrated components of the learning pathway. This creates a situation in which classroom instruction ends without a bridge to the contexts where communicative competence must ultimately function.

Benson (2011) argues that out-of-class language learning is not merely supplementary but constitutive of communicative competence; learners require autonomous, self-directed engagement with the language in authentic contexts to develop the flexibility and confidence that classroom interaction alone cannot produce. However, autonomy does not emerge spontaneously; it requires institutional scaffolding, structured task design, and accessible feedback mechanisms (Little, 2007).

In Vietnamese universities specifically, structural disconnection is exacerbated by large class size, limited contact hours, and the absence of systematic co-curricular English programming. Students may be aware that they need more practice but lack the institutional framework, the tasks, feedback channels, and social structures, that would make sustained out-of-class practice feasible and motivating.

Synthesis: The Non-Use Loop and the Case for an Integrated Intervention

These three themes are not independent; they interact to form a self-reinforcing cycle. Psychological inhibition reduces learners' willingness to engage in speaking practice. Methodological passivity means that even when students do study, they generate insufficient output to build confidence or automaticity. Structural disconnection ensures that no institutional mechanism intervenes to break the cycle. Together, these themes explain why many Vietnamese university students can comprehend English but cannot produce it fluently under communicative pressure. This thematic analysis thus establishes the rationale for a multi-level intervention that simultaneously addresses all three themes: psychological safety design

(Theme 1), output-driven task structures (Theme 2), and institutional alignment mechanisms (Theme 3). The solutions proposed in Chapter 4 (Solutions) are each mapped to these themes, ensuring that recommendations are grounded in the identified evidence base rather than general prescription.

Solutions

Individual Solution: Build a Micro Roadmap and Daily Output Routines

(Addresses Themes 1 and 2)

Rather than waiting for external conditions to emerge, students should begin with personal habits. The key approach is to design a micro-learning roadmap, a short but sustainable plan that students can maintain daily. Evidence from research on deliberate practice (Ericsson et al., 1993) and distributed learning (Dörnyei, 2005) consistently shows that frequent, low-intensity practice is more effective for skill development than infrequent, high-intensity study sessions.

A practical cycle can follow three steps grounded in Swain's (1985) input-output model: (1) receive comprehensible input at a slightly challenging but manageable level, (2) process the input to generate active language through output (speaking or writing), and (3) reflect on or record feedback to consolidate progress. For example, a 25–30-minute session may include listening to a short clip or discipline-related podcast for 8–10 minutes, followed by writing or recording a three-to-five-sentence summary, then practicing a one-minute spoken response. The crucial point is not how much content is studied, but that each session produces a language output product. With output appearing consistently, learners gradually overcome reticence, because they discover they can complete small tasks daily, and small repeated efforts build confidence over time.

To address Theme 1 directly, students should adjust their self-evaluation criteria. Research on WTC (MacIntyre et al., 1998) demonstrates that learners' communicative confidence is strongly shaped by their perception of success in prior speaking attempts. Adopting a “speak to practice” orientation, accepting that errors will occur and be corrected iteratively, is therefore not merely motivational advice but an evidence-based strategy for increasing WTC.

A helpful mechanism is keeping a voice journal using fixed sentence frames (e.g., “From what I heard…,” “The main point is that…,” “In my opinion, … because …”). Research on formulaic language (Wray, 2002) confirms that such frames reduce cognitive load, freeing attentional resources for meaning-making rather than grammatical monitoring.

Finally, motivation can be maintained through simple progress tracking. Dörnyei (2005) identifies self-efficacy, learners' belief in their own capacity to succeed, as one of the strongest predictors of sustained language learning effort. A brief daily learning diary (input received, output produced, errors noticed, next practice target) operationalizes this by making progress visible and directional.

Group/Community Solution: Create a “Safe Space for Speaking” with Task-Based Interaction

(Addresses Themes 1 and 2)

Lasting improvement in communicative competence typically requires interaction with others. Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD) and its application in collaborative language learning (Swain & Lapkin, 1998) provide theoretical

justification: learners develop more effectively when collaborating with peers who can provide support slightly beyond their current independent capacity.

Groups should operate with clear norms from the outset: (1) treat errors as data for learning rather than failures; (2) provide feedback as suggestions for improvement rather than evaluation; (3) rotate speaking roles to ensure equal participation; and (4) run activities with tasks and time limits. These norms directly address the fear of negative evaluation identified in Theme 1.

A suitable mechanism is the task-based learning cycle (Ellis, 2003): (1) warm-up with accessible questions, (2) task execution (group discussion, role-play, peer interviews, debates requiring reasoned justification), and (3) end-of-session output (a one-to-two-minute speaking segment or a short written paragraph). Critically, feedback should operate on two levels: first on meaning (whether the main point is clear), then on language (vocabulary, structures, and recurring errors). Lyster and Saito's (2010) meta-analysis of oral feedback in classroom SLA found that corrective feedback, particularly recasts and prompts, significantly promotes L2 development, with the greatest effects when feedback is explicit and focused on specific linguistic targets. Two-level feedback respects this evidence by prioritizing communication before correction.

A buddy system can further reduce anxiety. Because partners are familiar, social risk is reduced and speaking becomes more natural. Scripts aligned with academic and career needs (e.g., presenting a study project, describing research habits, introducing skills) ensure that practice is purposeful rather than merely social.

University Solution: Align Classroom Learning with Out-of-Class Practice

(Addresses Theme 3)

Universities must create structural alignment mechanisms that connect in-class activities with out-of-class practice, directly addressing Theme 3. Without institutional scaffolding, even motivated students lack the frameworks to make sustained extracurricular practice feasible.

First, universities should embed an input-task-feedback cycle in each course. Rather than assigning generic homework, instructors should assign output-oriented tasks directly aligned with classroom input: summarizing a reading in English, recording a one-minute presentation, writing a short counterargument, or participating in structured group discussions. Little's (2007) research on learner autonomy demonstrates that structured task design is essential for enabling self-directed learning; autonomy does not emerge spontaneously but must be scaffolded through clear task parameters and feedback.

Second, universities should provide accessible learning spaces and channels: flexible practice rooms, an academic library with appropriate materials, or an online platform for submitting speaking and writing outputs. Standardized rubrics assessing clarity of meaning, appropriateness of vocabulary, and improvement over time reframe evaluation as developmental rather than purely summative, a shift supported by Hattie and Timperley's (2007) influential meta-analysis demonstrating that feedback oriented toward learning goals (rather than grades) produces significantly greater achievement gains.

Third, periodic events with clearly defined language objectives, such as English Speaking Day, discipline-based academic club activities, and mid-semester mini-presentations, should follow a gradual ladder design: beginners can start with very brief tasks (30-second speaking), then progressively increase length and interactional complexity. This scaffolded

approach prevents newcomers from experiencing overwhelming demands that would reinforce Theme 1 barriers.

Finally, universities should implement mechanisms to recognize progress, such as incremental participation credit, certificates, or tracked "hours of purposeful use," to integrate out-of-class learning into the official curriculum rather than treating it as optional.

Community/Partner Solution: Extend English Output to Authentic Contexts

(Addresses Themes 1, 2, and 3)

To make English genuinely serve academic and professional development, learners must experience English in contexts approximating real-world use. This addresses all three themes simultaneously: authentic contexts reduce the artificial performance pressure of Theme 1, generate genuine communicative output beyond Theme 2's passive reception, and extend the structural reach of the institution beyond Theme 3's classroom walls.

Effective approaches involve activities with real communicative purposes: alumni career talks, mock interviews with professionals, and collaborative projects with organizations (writing service brochures, producing academic communication content, editing project proposals). These generate authentic output; students must communicate to achieve project goals, not merely to satisfy academic requirements. Research on task authenticity (Ellis, 2003) consistently shows that learners demonstrate greater effort, more complex language use, and higher motivation when tasks have purposes beyond the classroom.

Community settings also support informal interaction: language exchange sessions with international residents or visitors, and community seminars on local topics (environment, culture, history) with English discussion segments. When students recognize that they are speaking to address factual issues or share genuine opinions rather than to complete assignments, communicative motivation increases significantly, consistent with self-determination theory's emphasis on autonomy and purpose as drivers of intrinsic motivation (Deci & Ryan, 1985, as cited in Dörnyei, 2005).

To ensure quality, universities and partners should provide task frameworks in advance: topic, product criteria, time allocation, and feedback guidance. This structure prevents learners from facing overly demanding situations without support, especially for students who are psychologically hesitant to make mistakes. Meanwhile, evaluation should focus on progress rather than final performance only, encouraging participation from students at all proficiency levels.

Conclusion, Recommendations, and Implementation Roadmap

5.1. Conclusion

This paper has argued that the limitations in Vietnamese university students' English speaking skills are not primarily linguistic in origin. Through thematic analysis, three interlocking barriers were identified: psychological inhibition (Theme 1), methodological passivity (Theme 2), and structural disconnection (Theme 3). These themes interact to form a non-use loop in which reduced speaking leads to lower confidence, which further reduces speaking practice.

The study's primary contribution to the existing TESOL literature lies in three areas. First, while prior research has examined individual barriers (anxiety, input-output imbalance, institutional constraints) in isolation, this paper provides an integrated thematic analysis that demonstrates how these barriers mutually reinforce one another, and therefore why single-level

interventions have limited effectiveness. Second, the proposed multi-level solution framework is explicitly mapped to the identified themes, ensuring that each recommendation is evidence-grounded rather than prescriptive. Third, the study situates this framework specifically within the Vietnamese university context, a setting characterized by large class sizes, grammar-translation traditions, face-saving cultural norms, and limited co-curricular infrastructure, thereby filling a gap in contextually specific intervention design for Southeast Asian EFL learners.

The theoretical foundation of this paper, drawing on Swain's Output Hypothesis, MacIntyre et al.'s WTC model, Long's Interaction Hypothesis, and Ellis's task-based learning framework, supports the view that effective out-of-class learning environments must coordinate exposure, output tasks, psychological safety, and feedback mechanisms. Future research should empirically evaluate the hypotheses proposed in Chapter 1 (Introduction) through quasi-experimental designs, examining whether the multi-level intervention framework produces measurable improvements in speaking proficiency and WTC compared to control conditions.

Recommendations by Stakeholder Group

Recommendations for Students (Individual Level)

- Design a micro-learning roadmap: each learning session (20–30 minutes) should produce at least one speaking/writing output (1–5 sentences, or a 30–60 second recording), consistent with output-driven SLA principles (Swain, 1985).
- Shift the goal from “speaking perfectly” to “speaking to practice,” prioritizing communicative intent, supported by WTC research demonstrating that perceived competence, not actual accuracy, most strongly predicts speaking engagement (MacIntyre et al., 1998).
- Use sentence frames to reduce cognitive load and speaking anxiety (Wray, 2002).
- Track progress through behavioral indicators: number of recordings/speaking attempts, frequency of repeated errors, and improvement in clarity of meaning.

Recommendations for Learning Groups/Clubs (Group Level)

- Design activities using the task-based learning cycle: warm-up → task execution → end-of-session output (Ellis, 2003).
- Establish two-layer feedback: content clarity first, then language accuracy, consistent with Lyster and Saito's (2010) finding that focused, staged feedback maximizes uptake without increasing anxiety.
- Apply a buddy system and rotate speaking roles to ensure equal practice opportunities for all members.

Recommendations for Universities (Institutional Level)

- Embed an out-of-class practice cycle in each course: in-class input → output-based tasks → product submission → rubric-based feedback (Hattie & Timperley, 2007).
- Build accessible submission and feedback channels to reduce the psychological cost of evaluation (Horwitz et al., 1986).

- Organize periodic speaking events (e.g., speaking day, mini presentations, academic-club seminars) with a gradual ladder design, brief, low-stakes tasks first, increasing in length and complexity (Skehan & Foster, 1999).

Recommendations for Communities and Partners (External Level)

- Create activities with authentic communicative purpose (career talks, mock interviews, academic communication projects) (Ellis, 2003).
- Provide clear task frameworks in advance (time limits, expected products, feedback criteria).
- Base evaluation on progress, not only final performance, to encourage broad participation.

Sample Implementation Roadmap (8–10 Weeks)

The following roadmap is designed for a club, speaking-support module, or group intervention program. It is organized around the three identified themes and follows the principle: psychological safety (Theme 1) + output tasks (Theme 2) + structural connection (Theme 3) + gradual extension to authentic contexts.

Week 1: Establish Psychological Safety and Speaking Norms (Theme 1)

- Individual: begin a voice journal (2–3 times/week), 30–45 seconds using sentence frames.
- Group: agree on feedback norms; run a "round-robin" activity (30 seconds per learner on a familiar topic).
- Output: one personal recording plus two-layer peer feedback.

Week 2: Build Basic Output Chains (Theme 2)

- Individual: complete three "listen → summarize (3–5 sentences) → speak (60 seconds)" cycles.
- Group: implement a buddy system with scripted Q&A (self-introduction, learning goals).
- Output: each pair submits a short, recorded dialogue.

Week 3: Increase Task Complexity (Theme 2)

- Group: discussion tasks with a "claim-reason-example" structure.
- Feedback: rubric focusing on two main criteria, clarity of meaning and appropriateness of vocabulary.
- Output: one-to-two-minute speaking segment per group.

Week 4: Integrate Course Content into Speaking Tasks (Theme 3)

- University/Instructor: assign speaking output tasks aligned with in-class content.
- Individual: record a one-to-two-minute talk and write an 80–120-word summary.
- Output: products submitted for rubric-based feedback.

Week 5: Develop Fluency through Re-telling (Themes 1 and 2)

- Individual: re-telling practice, listen to a previous clip and retell it with targeted improvement (1–2 aspects).
- Group: role-play academic situations (challenging an idea, reporting project progress).
- Output: Version 1 and Version 2 re-tellings within the same week.

Week 6: Move to Academic-Style Speaking (Theme 2)

- Group: mini-debate or seminar-style discussion with assigned roles.
- Feedback: introduce an "error log" to identify and target recurring group-level errors.
- Output: two-to-three-minute presentation or debate segment.

Week 7: Connect to Authentic Contexts (Themes 3 and 4)

- Invite an alumni speaker or professional staff member for a structured Q&A in English.
- Learners complete mock interviews based on a provided question set.
- Output: recorded answers plus a brief self-reflection on improvement.

Week 8: Consolidate and Evaluate Progress (Themes 1–3)

- Individual: produce a "before-after" reflection comparing early and final recordings.
- Group: deliver a consolidated topic presentation; receive rubric-based feedback.
- Final output: three-to-five-minute speaking performance plus a four-week self-practice plan.

If extending to 10 weeks, add two weeks for "guided freer speaking tasks" and a one-peer-coaching cycle between groups.

In short, the core strategy is to engineer an environment and routine in which speaking becomes a repeated, supported, and purpose-driven practice.

References

Benson, P. (2011). *Teaching and researching autonomy in language learning* (2nd ed.). Pearson Longman.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541–577. https://doi.org/10.1017/S0272263102004023

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3*(2), 73–93. https://doi.org/10.1080/09658416.1994.9959845

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Little, D. (2007). Language learner autonomy: Some fundamental considerations revisited. *Innovation in Language Learning and Teaching, 1*(1), 14–29. https://doi.org/10.2167/illt040.0

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*(2), 126–141. https://doi.org/10.1093/applin/4.2.126

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition, 32*(1), 1–31. https://doi.org/10.1017/S0272263109990520

MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching, 7*(1), 61–88. https://doi.org/10.14746/ssllt.2017.7.1.4

Nguyen, H. T. M. (2011). Primary English language education policy in Vietnam: Insights from implementation. *Current Issues in Language Planning, 12*(2), 225–249. https://doi.org/10.1080/14664208.2011.597048

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93–120. https://doi.org/10.1111/0023-8333.00039

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford University Press.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320–337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex Publishing.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Soberman, Eds.). Harvard University Press.

Wray, A. (2002). *Formulaic language and lexicon*. Cambridge University Press.

#105 EXPLORING FIRST-YEAR ENGLISH-MAJORED STUDENTS' PERCEPTIONS OF PEDAGOGICAL TRANSLANGUAGING IN AN EMI CONTEXT

Thinh Tien Vu[1], Cam Diem Nhi Hoang[2]

Abstract

The rapid expansion of English as a Medium of Instruction (EMI) has highlighted the complex roles of the native language (L1) and the target language (L2) in academic learning. Pedagogical translanguaging has emerged as a strategic method to support student comprehension within these environments. This study investigated how English-major students perceived teachers' pedagogical translanguaging practices and identified the specific practices most and least valued in EMI classrooms. The participants included fifty first-year English-majored students at a public university in Vietnam. Employing a quantitative research design, the data were collected via a structured questionnaire and analysed using descriptive statistics. The results indicated that students generally hold positive perceptions of pedagogical translanguaging across three primary dimensions: Stance, Design, and Shift. Furthermore, the three most frequently reported practices involve using Vietnamese to clarify difficult concepts, explaining complex academic content bilingually, and supporting vocabulary and grammar acquisition through both languages. These findings suggest that pedagogical translanguaging is perceived as a useful instructional approach that helps to facilitate learning in EMI contexts.

Keywords: English Medium of Instruction (EMI), English-majored Students, Pedagogical Translanguaging, Student Perceptions, Translanguaging practices

1. Introduction

English holds an unparalleled position in global internationalisation and has increasingly been adopted as the medium of instruction in educational settings (Liu, 2021). In recent years, English as a Medium of Instruction (EMI) at universities has been significantly growing as an educational phenomenon. It is increasingly attracting researchers' attention, particularly those in applied linguistics (Macaro, 2019). Alongside this development, for decades, foreign language teaching has been guided by the principle that teachers should use only the target language (TL) and avoid the native language (L1) unless necessary (Littlewood & Yu, 2009). Due to ongoing debates surrounding bilingual language use in EMI classrooms, a new and developing teaching and learning method called Translanguaging, especially Pedagogical Translanguaging, has emerged as a framework that reconceptualises the use of the native language (L1) in the foreign language classroom by arguing that the use of L1 alongside the target language is considered a functional language practice (Sapitri et al., 2018).

Translanguaging has emerged as a response to EMI programs for non-native English learners, addressing challenges such as monolingual bias and the difficulties faced by learners

[1] *Corresponding author, PhD candidate, School of Languages, International University - VNU HCM, HCMC, Vietnam;* vtthinh@hcmiu.edu.vn
[2] *BA, School of Languages, International University - VNU HCM, HCMC, Vietnam*

with low English proficiency (Nguyen, 2022). According to Fang et al. (2022), Translanguaging, in both practice and pedagogy, has attracted growing interest for its recognition of the linguistic resources and cultural capital of both teachers and learners.

A study by Zumor (2019) conducted in an EMI environment found that the language barrier was a significant factor in student failure. This language barrier creates an immediate academic crisis for Vietnamese first-year English-majored students, who are suddenly forced to shift to a completely English-speaking environment. In practice, Iversen (2019) stated that students with low-level language proficiency often use translanguaging as a supplementary tool to understand the content of a text or lesson. This led to the emergence of translanguaging as the most intuitive and optimal problem-solving method, given its benefits. The use of translanguaging promotes students' metacognitive and critical-thinking abilities, while providing cognitive support (García & Wei, 2014). There is a significant gap in empirical and contextual research, as previous literature has focused primarily on teachers' perspectives, completely neglecting quantitative data on the perceptions of first-year English-majored students regarding the use of pedagogical translanguaging in EMI environments within a Vietnamese context. This indicates that research on students' perceptions of pedagogical translanguaging is extremely useful at this stage.

This research explored first-year English-majored students' perceptions of pedagogical translanguaging in an EMI context and identified the practices most and least valued by students in EMI classrooms. Other student groups (second-year to fourth-year students) were intentionally excluded from this study because first-year students face the greatest language shock during the first semester of the transition to English-language instruction at the university level. Theoretically, it not only contributed to the limited body of research on pedagogical translanguaging in teaching and learning within EMI environments but also shifted the focus from teachers to students. In practical terms, this study offers a deeper understanding of learners' experiences of pedagogical translanguaging in the EMI context. Therefore, teachers and curriculum designers can know how students engage with translanguaging activities, thereby supporting more effective pedagogical decisions in EMI classrooms. To fulfil the purpose of the study, the paper was conducted to answer the following research questions:

1. How do first-year English-majored students perceive teachers' pedagogical translanguaging practices in EMI classes?
2. Which pedagogical translanguaging practices are most and least valued by first-year English-majored students in EMI classrooms?

2. Literature Review

2.1. EMI Context

EMI has emerged as a rapidly expanding global phenomenon, with its demand spreading worldwide only over the past two decades (Romanowski, 2020). Macaro (2019) defined EMI (English Medium of Instruction) as the provision of instruction in contexts where English is not the commonly used language. Dearden (2014) suggested EMI as the implementation of using English to teach academic subjects in countries or jurisdictions where the mother tongue (L1) of most of the population is not English. In this study, EMI is the use of English to teach academic subjects in countries such as Vietnam, where the majority of the population speaks a language other than English.

Beyond promoting internationalisation and globalisation, EMI has been shown to benefit students whose first language is not English, as it helps develop their English proficiency, prepares them for the job market, and enables them to access abundant academic resources in English to succeed in higher education (Almatrafi, 2023).

In Vietnam, EMI programs may be implemented with MOET approval when institutions meet requirements related to teaching staff, facilities, and resources, or when they collaborate with qualified international partners (Vu et al., 2025).

2.2. Perception

In the field of education, André et al. (2020) argued that students' perceptions of actual classroom activities can be more important than what outsiders observe, because students' perceptions guide their own learning behaviour, based on their own understanding. In this study, perception refers to the beliefs or opinions held by first-year English-majored students, based on how things seem.

2.3. Translanguaging

The concept of translanguaging is understood as a method of using language to integrate other linguistic resources, guided by pedagogical principles, to fully exploit students' multilingual abilities (Romanowski, 2020). In general, translanguaging refers to allowing students to draw on their native language repertoires to support learning in a target-language classroom (Khairunnisa & Lukmana, 2020). According to Nguyen (2022), translanguaging is viewed from two perspectives: a sociolinguistic one, which describes the fluid language practices of communities, and a pedagogical one, where teachers and students collaboratively create spaces that embrace diverse language use to support creative and critical learning.

In the context of education, translanguaging refers to the strategic and purposeful use of learners' language by teachers to support the learning process by leveraging learners' full linguistic resources, which are collectively stored (Barahona, 2020).

Translanguaging, beyond educational contexts, is mentioned by Sahib (2019), who claims that it has occurred outside the classroom. However, in this study, the focus is on the educational context, thereby omitting spontaneous translanguaging.

2.4 Translanguaging and Code-Switching

Sahib (2019) claimed that translanguaging and code-switching share certain similarities, such as both involving the use of students' first language to support second-language learning in the classroom and both entailing shifting between languages during communication. However, Cahyani et al. (2016) clarified that translanguaging differs from code-switching in that it is not simply a conversion between two separate monolingual languages but rather combines the two languages into a unified whole to achieve effective communication.

This study focuses solely on translanguaging practices, as they provide a more comprehensive and holistic framework for understanding multilingual meaning-making in EMI classrooms.

2.5. Pedagogical Translanguaging

Translanguaging pedagogy has increasingly attracted attention as it facilitates teaching and learning while allowing both learners and teachers greater flexibility in their language use for expression (Romanowski, 2020). Moraczewska (2024) stated that translanguaging pedagogy intentionally draws on students' existing linguistic resources to support knowledge construction and active participation in meaning-making.

Specifically, in this study, pedagogical translanguaging refers to the intentional classroom practice in which a teacher actively values students' full linguistic repertoires as a fundamental right, designs lessons that purposefully connect home languages to ensure complete comprehension and flexibly utilises spontaneous language shifts during interactions to support meaningful learning.

Cenoz and Gorter (2021) stated some characteristics of pedagogical translanguaging. First, this approach offers a wide range of practices and activities that draw on elements from the entire linguistic repertoire. It has different functions, such as gaining a deeper understanding of academic texts, accessing more information, developing metalinguistic awareness, and reinforcing skills through coordinated activities across different languages. Second, it can be applied to more programmes and a wide range of students at all levels of education (Cenoz & Gorter, 2021).

2.5.1. Benefits of Pedagogical Translanguaging

In the educational context, translanguaging is viewed as a scaffolding strategy but a holistic meaning-making process of language and learning (Pantalleresco, 2021). Baker and Wright (2011) emphasised four educational advantages of translanguaging in the classroom. First, he asserted that translanguaging helps learners better understand the subject matter through interdependence and language transfer between their native language (L1) and the target language (TL). Baker and Wright claimed that L1 offers learners greater comfort and a sense of safety. It is the strongest language, which can help develop skills in TL. These two benefits align well with the first dimension of the framework of García et al. (2017), which focuses on the Teacher's Stance. The third advantage mentioned by Baker and Wright (2011) is that multilingualism facilitates communication between school and family, especially when parents do not understand the target language their child is being taught. Finally, integrating both languages in the classroom allows learners to develop their second-language skills while learning the subject matter.

2.5.2. Disadvantages of Pedagogical Translanguaging

Allard (2017) found that participants had negative attitudes toward translanguaging because it distracted them and reduced their ability to absorb knowledge in the target language, leading them to prefer English-language instruction.

The overuse of the mother tongue in the classroom can hinder students' language development, as they have not been adequately exposed to the target language (Yuvayapan, 2019).

Furthermore, translanguaging can lead to some misunderstandings between languages, as students may misuse vocabulary by relying on inappropriate Vietnamese equivalents (Bui & Nguyen, 2014).

2.5.3. Translanguaging Practices

Regarding pedagogical translanguaging practices, Cenoz and Gorter (2021) stated that they aimed to activate prior multilingual knowledge in different forms. There are five

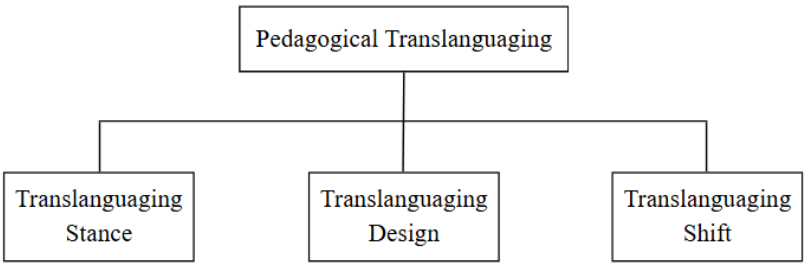
pedagogical purposes in this study: explaining concepts, checking comprehension, localising content, reinforcing instructions, and building relationships (Fang & Liu, 2020).

2.6. Conceptual framework

This study employed a pedagogical framework of García et al. (2017), which formulated three components: translanguaging Stance, translanguaging Design, and translanguaging Shift.

Figure 1

Pedagogical Translanguaging framework by García et al. (2017)



To be more specific, the Stance refers to "teacher's belief that a bilingual student has one holistic language repertoire that he or she draws on at school" (Turnbull, 2017, p.1). The Design involves the deliberate integration of students' native and school languages into classroom activities, while the Shift represents the emergence of unplanned, spontaneous translanguaging, with the teacher making immediate decisions that demonstrate flexibility and a willingness to support students' voices (Moraczewska, 2024).

In general, these three dimensions will be used in this study to address the two research questions mentioned above.

2.7. Previous studies

Several previous studies have explored subjects' perceptions of translanguaging in various countries worldwide, including students and teachers. Romanowski (2020) conducted research to explore students' and teachers' perceptions of translanguaging in EMI classrooms in Poland. Almatrafi (2023) conducted a study of Saudi EFL students to investigate their perceptions of pedagogical translanguaging in English-Medium Instruction. Both studies found that teachers and students had a positive outlook on classroom translanguaging. However, they mostly focus on teachers, who actively apply translanguaging in their teaching, neglecting in-depth research on students, who are the direct recipients of this practice.

In the Vietnamese context, the use of translanguaging has recently attracted researchers' attention; however, it remains understudied. A study conducted by Pham and Vu (2023) examined EFL secondary school teachers in both public and private schools in Hanoi. With a convergent mixed-methods study, the results showed that most teachers had positive attitudes toward translanguaging. They acknowledged its role in scaffolding students' learning, particularly for cognitive functions such as explaining grammar and vocabulary. Moreover, Ngo (2025) conducted research on EFL Learners' Perceptions of translanguaging in English classes at an HCM English Centre and found that students with lower levels of competence are more receptive to its use.

However, there is still little research on EMI in the university setting, where students' perspectives are considered important. In general, very few studies have sought to gather perspectives from linguistics students on how they perceive the use of translanguaging at the university level. These shortcomings in previous studies, along with the lack of research in the Vietnamese context, serve as inspiration for this study, which explores first-year English-majored students' perceptions of pedagogical translanguaging in an EMI context. This study is significant to conduct to enrich the literature on translanguaging.

3. Methodology

3.1. Study design

The study employs a quantitative design to investigate students' perceptions of teachers' translanguaging practices in EMI classrooms. Rahman (2016) noted that an advantage of quantitative research is that quantitative results can potentially be generalised to the entire population or to a smaller population group.

3.2. Participants

The study sample consisted of 50 first-year English Linguistics and Literature students at an EMI university in Ho Chi Minh City. The participants were recent high school graduates who were gradually adapting to a teaching and learning environment conducted entirely in English. Therefore, many first-year students may feel overwhelmed by the amount of information they receive in English.

Altheyab and Alalwi (2022) reported that EMI implementation can yield disappointing results, especially for students whose English proficiency is not yet sufficient. More specifically, linguistics students were selected because the research aimed to understand their perceptions of teachers' pedagogical translanguaging practices in light of the changing language-learning environment.

The study was conducted after first-year students of the 2025-2026 cohort had completed their first semester. In their first semester of university, students who had spent sufficient time in an EMI environment were able to recognise the pedagogical translanguaging techniques instructors used to support learning.

3.3. Sampling

This study utilises a non-probability purposive sampling method to ensure that all participants are first-year English-majored students experiencing the phenomenon under study. The purposive sampling technique involves selecting candidates across a broad spectrum relating to the topic of study (Etikan, 2016). The primary goal of targeted sampling is to focus on specific characteristics of the population of interest, thereby increasing the possibility of collecting the desired data (Rai & Thapa, 2015).

In this study, participants had to be first-year English majors studying in an EMI environment with pedagogical translanguaging applied by the instructors.

3.4. Instruments

A structured questionnaire was designed to collect data on first-year English majors' perceptions of pedagogical translanguaging in EMI classrooms. The questionnaire is based on the theoretical framework of pedagogical translanguaging proposed by García et al. (2017), encompassing three components: Stance, Design, and Shifts.

The first section gathered demographic information about participants. The second section included 15 Likert-scale items rated on a 5-point scale (1 = Strongly Disagree, 5 = Strongly Agree) regarding students' perceptions of the three dimensions mentioned above. To be specific, items B1-B3 were used to explore student perceptions of Translanguaging Practices – dimension Stance; items B4-B9 were used to explore the dimension Design; and items B10-B15 were used to examine the dimension Shift. The third section comprised 3 questions, which aimed to explore overall perceptions of teachers' pedagogical translanguaging.

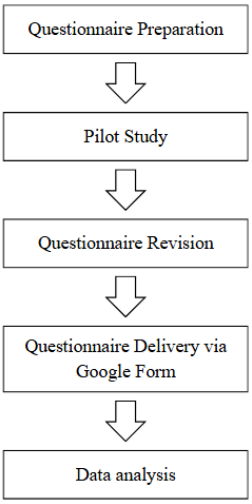
The questionnaires were conducted via Google Forms because it enables reaching many participants within a limited time. It ensures anonymity, encouraging more honest and objective responses. Besides, the digital format minimises data-entry errors and facilitates automatic sorting and export to SPSS for statistical analysis. Additionally, Google Forms is more cost-effective and accessible, as students can conveniently complete the questionnaire using their personal devices.

3.5. Data collection

As illustrated in Figure 2, the data collection process began with selecting a research topic, followed by the development of a structured questionnaire and a study outline. The questionnaire was based on the conceptual framework grounded in the García et al. (2017) study, which includes three dimensions: Stance, Design, and Shift.

Figure 2

Data collection procedure



A pilot survey was conducted to assess the clarity, validity, and reliability of the instrument. The questionnaire was pilot-tested with 16 first-year English-majored students and demonstrated strong internal consistency and acceptable reliability (Cronbach's alpha = .93). The questionnaires were conducted via Google Forms, which were cost-saving and easy to analyse. Pilot feedback led to minor changes in question wording to improve clarity. For

example, the previous statements I wrote: "My teachers *show openness* to students using Vietnamese to clarify or negotiate meaning." The pilot respondents hesitated to rate this item because of the phrase "show openness." Therefore, I changed it to "My teachers *allow students to use Vietnamese* to clarify or negotiate meaning," which helps the statement be clearer for participants to rate.

For the main data collection stage, a consent form was included at the beginning, outlining participants' rights and ethical considerations. After collecting responses, the internal consistency of the main dataset was evaluated using Cronbach's alpha coefficients to ensure instrument reliability. The overall questionnaire demonstrated high reliability with an alpha value of .95, while the individual subscales for Stance, Design, and Shift yielded acceptable internal consistency values of .83, .89, and .83, respectively. These values suggested that the adapted instrument was reliable for further descriptive statistical analysis. After reliability was checked, the researcher conducted a descriptive statistical analysis. Sloman (2010) stated that the most used descriptive analysis approach for research purposes involves recording behaviour and environmental events through continuous or interval recording procedures. Therefore, this technique is appropriate in this research.

3.6. Data analysis

After completing the online survey, responses were exported directly from Google Forms to Google Sheets. The dataset was then transferred to Microsoft Excel for initial screening, during which inappropriate submissions were checked and invalid responses removed to ensure data reliability.

The data were analysed using IBM SPSS Statistics Version 26 (SPSS) for quantitative analysis. The mean (M) and standard deviation (SD) were calculated for each item. SPSS was chosen because it is widely used in quantitative research in applied linguistics and education.

Descriptive statistics were used to summarise participants' perceptions of pedagogical translanguaging. This procedure may address how students perceive teachers' PT practices and which PT practices are most frequently used in EMI classrooms.

4. Results

4.1. Students' perceptions of pedagogical translanguaging in EMI classes

The descriptive statistical analysis indicates a positive tendency toward pedagogical translanguaging in EMI classes. The overall mean for the 21 items was 4.16 (SD = 0.88), suggesting that the strategic use of Vietnamese alongside English is a beneficial instructional approach.

Table 1 presents the four mean scores for pedagogical implementation.

Table 22

Students' perceptions of pedagogical translanguaging in EMI classes (N = 50)

Dimension	Mean	SD
Stance	4.0	0.91

Design	4.20	0.91
Shift	4.24	0.84
Overall	4.20	0.86
Total	4.16	0.88

The Shift dimension received the highest mean score (M = 4.24), followed closely by Design (M = 4.20), indicating that the implementation of pedagogical translanguaging in EMI classrooms is an appreciably recognised and valued practice among the participants.

4.2. Overall students' perceptions

Table 2 below shows a general view of students' perceptions of pedagogical translanguaging in EMI classes. The overview of students' perceptions of pedagogical translanguaging in EMI classrooms was evaluated using three main items (C1-C3). The overall mean score was 4.20 (SD = 0.86), indicating that students responded positively to the implementation of pedagogical translanguaging in EMI classrooms.

Table 23.

Overall perceptions

No.	Item	Mean	SD
C1	Pedagogical translanguaging helps me learn more effectively in EMI classes	4.24	0.92
C2	Pedagogical translanguaging has had a positive impact on my learning experience in EMI classes	4.18	0.80
C3	Pedagogical translanguaging practices should be applied in future EMI classes	4.18	0.87
	Overall Mean	4.2	0.86

Both C2 and C3 received a mean score of 4.18. These statistics show that students view pedagogical translanguaging as useful to their learning. Item C1 has the highest mean of 4.24 (SD = 0.92), indicating that first-year students consider pedagogical translanguaging a helpful approach in EMI classrooms.

4.3. Perceptions of the teacher's Stance

Regarding the first dimension, Stance (Table 3), the data collected from questions B1-B3 reveal most respondents' perception that their teachers positively support the use of Vietnamese in EMI classes.

Item B1 received the highest mean score of 4.20 (SD = 0.81), indicating that many students reported that their teachers allow the use of Vietnamese to enhance understanding and clarify meaning during lessons. Moreover, respondents agreed that their teachers view the use

of Vietnamese in class as part of the EMI learning environment, emphasising the vital role of pedagogical translanguaging in the classroom.

Table 3.

Perceptions of the Teacher's Stance

No.	Item	Mean	SD
B1	My teachers respond positively when Vietnamese is used to support understanding in class	4.20	0.81
B2	My teachers allow students to use Vietnamese to clarify negotiating meaning	3.98	0.94
B3	My teachers view the use of Vietnamese in class as part of the EMI learning environment	3.84	0.99
	Overall Mean	4.01	0.91

4.4. Perceptions of the teacher's Design

The second dimension concerned the teacher's Design of pedagogical translanguaging, as shown in Table 4. The teacher's Design was measured using six items (B4-B9), each reflecting a different aspect of learners' perceptions of teachers' pedagogical translanguaging in lessons.

Table 24.

Perceptions of the Teacher's Design

No.	Item	Mean	SD
B4	It helps me learn better when my teacher intentionally uses both Vietnamese and English to explain complex academic content.	4.48	0.68
B5	It helps me learn better when my teacher designs activities that allow discussion in both Vietnamese and English.	4.22	0.93
B6	It helps me learn better when my teacher encourages group work using both Vietnamese and English.	4.14	0.7
B7	It helps me learn better when my teacher allows students to ask and answer questions using both Vietnamese and English.	4.12	0.94
B8	It helps me learn better when my teacher uses Vietnamese to support understanding of materials in English (slides, textbooks, tasks).	4.14	1.03
B9	It helps me learn better when my teacher integrates students' life experiences into lessons through both Vietnamese and English.	4.08	0.92
	Overall Mean	4.19	0.89

Item B4 had the highest mean score of 4.48 (SD = 0.68). This showed that students strongly agreed that intentional bilingual explanation effectively facilitated their understanding of complex academic content, and the low standard deviation indicated that this perception was consistent across participants. This was followed by B5 (M = 4.22), B6 and B8 (M = 4.14), and B7 (M = 4.12), suggesting that implementing a multilingual approach in EMI classes helps students better understand the lessons.

In contrast, item B9 has the lowest mean at 4.08 (SD = 0.92), revealing that some students may view the integration of personal life experiences as less relevant to their learning.

4.5. Perceptions of the teacher's Shift

Regarding students' perceptions of the teacher's shift from item B10 to B15, as shown in Table 5, the statistics showed high overall scores, indicating that first-year students valued the spontaneous shift between English and Vietnamese during classes.

Item B10 had the highest mean of 4.50 (SD = 0.68), indicating that students learn better when their teacher uses Vietnamese to clarify difficult concepts in lessons. Additionally, item B13 has a high mean of 4.38 (SD = 0.86), indicating that students strongly value the use of both Vietnamese and English to explain unfamiliar vocabulary and grammar. This was followed by B13 (M = 4.38), B14 (M = 4.26), B12 (M = 4.16), and B11 (M = 4.10), suggesting that the flexibility between using the target language and the native language enhances students' comprehension and creates a safer environment for first-year students.

In contrast, item B15 had the lowest mean at 4.06 (SD = 0.98), suggesting that some students may disagree with the use of PT for their learning.

Table 25.

Perceptions of the teacher's Shift

No.	Item	Mean	SD
B10	It helps me learn better when my teacher uses Vietnamese to clarify difficult concepts.	4.50	0.68
B11	It helps me learn better when my teacher allows Vietnamese support while I try to answer in English.	4.10	0.95
B12	It helps me learn better when my teacher adjusts language use depending on students' proficiency levels.	4.16	0.77
B13	It helps me learn better when my teacher explains unfamiliar vocabulary or grammar using both Vietnamese and English.	4.38	0.86
B14	It helps me learn better when my teacher allows spontaneous discussions in Vietnamese and English during class to ensure complete comprehension.	4.26	0.83
B15	It helps me learn better when my teacher responds flexibly to students' questions by using both Vietnamese and English.	4.06	0.98
	Overall Mean	4.24	0.84

5. Discussion

5.1. RQ1: How do first-year English-majored students perceive teachers' pedagogical translanguaging practices in EMI classes?

Results from the questionnaire revealed that first-year English-majored students hold a positive perception of pedagogical translanguaging in the EMI context. The overall mean score of 4.14 indicates that integrating Vietnamese and English is a supportive tool that enhances students' understanding and fosters a supportive learning environment.

5.1.1 Perceived Teacher's Stance

Regarding the stance dimension, many respondents revealed that their teachers responded positively when Vietnamese was used to support comprehension in class (M = 4.20). This finding aligns with Pantalleresco (2021)'s research, which showed that teachers' acknowledgement of learners' L1 facilitates understanding and fosters a comfortable classroom environment. The teachers' view of Vietnamese as an essential component of the EMI learning environment received a high mean score (M=3.84), highlighting a shift toward the "meaning-making" process, in which L1 serves as a crucial tool for facilitating learning and understanding of the target language within a specific context (Liu, 2021). This practice helps narrow the gap between high- and low-proficiency students and first-year students, especially those with limited English proficiency, ensuring that all learners are accessible to the target knowledge (Pantalleresco, 2021).

5.1.2. Perceived Teacher's Design

Within the design dimension, participants perceived the teachers' intentional use of both Vietnamese and English to explain complex academic content as an effective scaffolding strategy that facilitates comprehension during cognitively demanding tasks. This result aligns with the suggestion by Alasmari et al. (2022) that using the mother tongue reinforces learning by clarifying complex concepts and terminology. Furthermore, many students expressed interest in activities that allowed discussion in both Vietnamese and English (M = 4.22). This result supports Pantalleresco's (2021) view that L1 not only facilitates meaning negotiation but also enhances the transition to the target language.

In contrast, some first-year students with high English proficiency expressed different views about using Vietnamese to aid understanding of English learning materials, such as slides, textbooks, and exercises (M = 4.14, SD = 1.03). This difference may be due to the overuse of L1 in the EMI environment, as noted by Phuong and Dang (2022). They found that over-reliance on the mother tongue can reduce learners' motivation to learn. While the mother tongue is a valuable pedagogical tool, its effectiveness depends more on the appropriate context used than on uncontrolled use (Phuong & Dang, 2022).

5.1.3. Perceived Teacher's Shift

Regarding the shift dimension, the collected data show that most learners agreed with teachers' use of Vietnamese to clarify difficult concepts (M = 4.50). This result aligns with the research of Liu et al. (2022), which found that translanguaging activities help create a sense of "safety" in the classroom, reducing learner anxiety and promoting collaborative learning. In the same vein, many participants agreed that using both Vietnamese and English to explain unfamiliar vocabulary and grammatical structures helps them learn better (M = 4.38). This could help address problems related to limited language proficiency in an EMI environment. The use of translanguaging in teaching can be seen as a bridge to close the language gap and

ensure equitable access to academic content. Almatrafi (2023) holds similar ideas, arguing that translanguaging helps lower-level learners feel more confident in participating and fully utilising their language skills.

However, not all participants supported teachers' flexible language switching. In contrast, some students reported that teachers' language switching could distract from their acquisition of the target language. This perceptual phenomenon may be due to differences in language proficiency among freshmen. Furthermore, the collected data indicate a negative view of allowing students to answer questions in both languages, which contrasts with findings from Nguyen's (2022) study, which argued that bilingualism improved communication effectiveness, reduced anxiety, and increased classroom engagement.

5.2. RQ2: Which pedagogical translanguaging practices are most and least valued by first-year English-majored students in EMI classrooms?

When ranked in descending order, the highest-rated practices were associated with the Shift aspect, followed by several selected practices within the Design aspect, all averaging above 4.3.

Specifically, the highest-rated item was B10 (*"It helps me learn better when my teacher uses Vietnamese to clarify difficult concepts")* (M = 4.50), which belongs to the Shift dimension. This was closely followed by B4 (*"It helps me learn better when my teacher intentionally uses both Vietnamese and English to explain complex academic content")* (M = 4.48), which belongs to Design dimension, and B13 *("It helps me learn better when my teacher explains unfamiliar vocabulary or grammar using both Vietnamese and English")* (M = 4.38), which belong to the Shift dimension. Overall, these highly ranked items highlight the crucial role of teachers' explanatory activities, particularly those that involve the flexible use of language within context to facilitate student comprehension.

Regarding item B10, in the context of EMI, where first-year students' English proficiency is diverse, students with lower proficiency are more likely to rely on their mother tongue (L1) to meet cognitive and linguistic demands in their studies. The importance of using L1 as a resource for comprehension in bilingual environments is supported by Cenoz and Etxague (2013). From that perspective, teachers' immediate use of Vietnamese serves as an essential scaffold to bridge the comprehension gap in real time.

Items B4 and B13 belong to the Design and Shift dimension. These two practices were identified by Nguyen et al. (2025) as the most common pedagogical functions of pedagogical translanguaging. Both B4 and B13 receive the highest mean scores because explanation plays a crucial role in ensuring students' comprehension in EMI. L1 not only enhances comprehension but also reduces cognitive load and anxiety in the classroom. This view is supported by As Yen and Danh (2022), who claimed that incorporating the native language can shorten explanation time and create a safe learning environment for students.

On the other hand, the two lowest-rated items were B2 and B3, which related to teachers' stance toward the use of Vietnamese in EMI classrooms. One possible explanation is that students enrolled in EMI classrooms often expect to study in a predominantly English environment. Some even associate this EMI environment with a learning condition like English immersion or study abroad programs, which helps to boost their language competence (Surtees, 2016). The above-mentioned viewpoint resonates with what Breeze and Roothooft (2021, p. 206) have found that some students might perceive the use of L1 as "dangerous territory",

reducing their exposure to English. This finding suggests that students may value a balance between maintaining the full English learning environment and strategically using Vietnamese for pedagogical purposes.

6. Conclusion and implications

This study sheds light on how first-year English-majored students perceived teachers' pedagogical translanguaging in EMI contexts. Overall, learners perceived pedagogical translanguaging as a useful tool to support comprehension. Additionally, students view the implementation of pedagogical translanguaging practices as a positive approach. Among the three dimensions (Stance, Design, Shift), the most valued translanguaging practices aligned with the "explaining concepts" pedagogical purpose.

The findings carry practical implications. The study can help teachers, schools, and educational institutions understand the potential of pedagogical translanguaging to support learning. Specifically, teachers will know how to design and implement lesson plans and activities suitable for students at varying levels. Schools and educational institutions can develop and deepen their understanding of pedagogical translanguaging.

Besides, the study has some limitations. A sample of 50 participants from one university may limit the generalizability of the findings, as it may not represent other educational settings in Vietnam. Besides, this study focuses solely on English majors, neglecting the perspectives of non-English majors. Future researchers should conduct research over longer periods, across diverse contexts, and with more varied participants to broaden the understanding of pedagogical translanguaging in EMI. Moreover, future studies should employ a mixed-methods approach rather than a quantitative approach alone to gain a deeper understanding of pedagogical translanguaging use in the context of EMI.

In conclusion, pedagogical translanguaging provides a promising approach to students' learning in an EMI environment. With appropriate instructional practices, this strategy has the potential to foster learners' comprehension, aligning with effective teaching and learning.

Acknowledgements

First and foremost, we would like to extend our thanks to International University, VNU HCM, and the School of Languages for providing us with all the necessary conditions to conduct this study. We would also like to extend our heartfelt appreciation to all the students who participated in this study. Your honest responses and the valuable time you contributed to this research played a significant role in its success. We are also deeply grateful to the teachers who kindly permitted us to enter their classrooms to collect data. We also sincerely thank the conference chairs and reviewers for their constructive feedback on earlier versions of this article.

References

Alasmari, M., Qasem, F., Ahmed, R., & Alrayes, M. (2022). Bilingual teachers' translanguaging practices and ideologies in online classrooms in Saudi Arabia. *Heliyon, 8 (9), e10537*. https://doi.org/10.1016/j.heliyon.2022.e10537

Allard, E. C. (2017). Re-examining teacher translanguaging: An ecological perspective. *Bilingual Research Journal*, *40*(2), 116-130. https://doi.org/10.1080/15235882.2017.1306597

Almatrafi, B. (2023). Saudi EFL Students' Perceptions of Pedagogical Translanguaging in English Medium Instruction. *International Journal of Language and Literary Studies*, *5*(4), 172-191. https://doi.org/10.36892/ijlls.v5i4.1460

Altheyab, N. A., & Alalwi, F. S. (2022). English-Medium Instruction and Content Learning in Freshman Year: An Investigation of a Saudi University Students' Challenges and Learning Strategies. *World Journal of English Language*, *12*(6), 71-82. https://doi.org/10.5430/wjel.v12n6p71

Alvi, M. (2016). *A manual for selecting sampling techniques in research*. Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/70218/

André, S., Maulana, R., Helms-Lorenz, M., Telli, S., Chun, S., Fernández-García, C.-M., De Jager, T., Irnidayanti, Y., Inda-Caro, M., Lee, O., Safrina, R., Coetzee, T., & Jeon, M. (2020). Student Perceptions in Measuring Teaching Behavior Across Six Countries: A Multi-Group Confirmatory Factor Analysis Approach to Measurement Invariance. *Frontiers in Psychology*, *11*, 1-19. https://doi.org/10.3389/fpsyg.2020.00273

Baker, C., & Wright, W. E. (2011). *Foundations of bilingual education and bilingualism*. Multilingual Matters Limited.

Barahona, M. (2020). The potential of translanguaging as a core teaching practice in an EFL context. *System*, *95*, 102368. https://doi.org/10.1016/j.system.2020.102368

Breeze, R., & Roothooft, H. (2021). Using the L1 in university-level EMI: Attitudes among lecturers in Spain. *Language awareness*, *30*(2), 195-215. https://doi.org/10.1080/09658416.2021.1903911

Bui, P. H., & Nguyen, T. T. A. (2014). The use of Vietnamese in English language classes-benefits and drawbacks. *International Journal on Studies in English Language and Literature*, *2*(12), 24-26.

Cahyani, H., Courcy, M. D., & Barnett, J. (2016). Teachers' code-switching in bilingual classrooms: Exploring pedagogical and sociocultural functions. *International Journal of Bilingual Education and Bilingualism*, *21*(4), 465-479. https://doi.org/10.1080/13670050.2016.1189509

Cenoz J, Gorter D (2021). *Pedagogical Translanguaging*. Cambridge University Press.

Cenoz, J., & Etxague, X. (2013). Spanish and English in Higher Education. *Bilingual and multilingual education in the 21st century: Building on experience*, *94*, 85-106.

Dearden, J. (2014). *English as a medium of instruction-a growing global phenomenon*. British Council.

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

Fang, F., Zhang, L. J., & Sah, P. K. (2022). *Translanguaging in Language Teaching and Learning: Current Practices and Future Directions*. *RELC Journal*, *53*(2), 305-312. https://doi.org/10.1177/00336882221114478

Fang, F., & Liu, Y. (2020). 'Using all English is not always meaningful': Stakeholders' perspectives on the use of and attitudes towards translanguaging at a Chinese university. *Lingua*, *247*, 102959. https://doi.org/10.1016/j.lingua.2020.102959

García, O., Johnson, S. I., & Seltzer, K. (2017). *The translanguaging classroom: Leveraging student bilingualism for learning*. Caslon.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave.

Iversen, J. Y. (2019). Pre-service teachers' translanguaging during field placement in multilingual, mainstream classrooms in Norway. *Language and Education*, *34*(1), 51-65. https://doi.org/10.1080/09500782.2019.1682599

Khairunnisa, K., & Lukmana, I. (2020). Teachers' Attitudes towards Translanguaging in Indonesian EFL Classrooms. *Jurnal Penelitian Pendidikan*, *20*(2), 254-266. https://doi.org/10.17509/jpp.v20i2.27046

Liu, D. (2021). A study of graduate students' perceptions towards pedagogical translanguaging at an international university in Bangkok. *Language in India*, *21*(3), 1-36. https://www.languageinindia.com/march2021/liutranslanguagingfinal.pdf

Liu, D., Deng, Y., & Wimpenny, K. (2022). Students' perceptions and experiences of translanguaging pedagogy in teaching English for academic purposes in China. *Teaching in Higher Education*, *29*(5), 1234-1252. https://doi.org/10.1080/13562517.2022.2129961

Littlewood, W., & Yu, B. (2009). First language and target language in the foreign language classroom. *Language Teaching*, *44*(1), 64-77. https://doi.org/10.1017/s0261444809990310

Macaro, E. (2019). Exploring the role of language in English medium instruction. International *Journal of Bilingual Education and Bilingualism, 23*(3), 263-276. https://doi.org/10.1080/13670050.2019.1620678

Moraczewska, A. (2024). Pedagogical Translanguaging for Epistemic Inclusion: Scaffolding Strategies in Bilingual Subject Instruction in Norway. *International Journal of Applied Linguistics, 36,* 39-50. https://doi.org/10.1111/ijal.12661

Ngo, K. T. (2025). EFL Learners' Perceptions of Translanguaging in English classes at an HCM English Centre. *International Journal of Language Instruction*, *4*(1), 50-64. https://doi.org/10.54855/ijli.25414

Nguyen, H. T., Nguyen, H. T., Gilanyi, L., Hoang, T. H., & Gao, X. (2025). Content Language Integrated Learning (CLIL): Teachers' metacognitive understanding of pedagogical translanguaging. *Learning and Instruction*, 97, 102085. https://doi.org/10.1016/j.learninstruc.2025.102085

Nguyen, T. N. T. (2022). A Review of Studies on EFL Teachers' and Students' Perceptions of Tranglanguaging as a Pedagogical Approach. *International Journal of TESOL & Education*, *2*(3), 324-331. https://doi.org/10.54855/ijte.222322

Pantalleresco, J. (2021). *The teachers' beliefs on translanguaging practices during speaking tasks in the Spanish as a foreign language classroom* (Master's dissertation, University of Malta). University of Malta Institutional Repository. https://www.um.edu.mt/library/oar/handle/123456789/102446

Phuong, H. Y., & Dang, T. D. (2022). English - majored students' perceptions on teachers' use of Vietnamese in EFL classrooms. *Journal of English Education and Linguistics Studies (JEELS)*, *8*(2), 371-391. https://doi.org/10.30762/jeels.v8i2.3379

Rahman, M. S. (2016). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: a literature review. *Journal of Education and Learning,* *6*(1), 102. https://doi.org/10.5539/jel.v6n1p102

Romanowski, P. (2020). Perceptions of translanguaging among the students and teachers of EMI classrooms in Poland. *Lenguas Modernas*, (55), 151-165.

Sahib, R. B. (2019). The use of translanguaging as a pedagogical strategy in efl classroom: A case study at bulukumba regency. *LET: Linguistics, Literature and English Teaching Journal*, *9*(2), 154-180. https://doi.org/10.18592/let.v9i2.3124

Sapitri, N. M., Batan, I. G., & Myartawan, I. P. N. W. (2018). Functions of teachers' translanguaging in the EFL classroom at two junior high schools in Singaraja. *Lingua Scientia, 25*(1), 29. https://doi.org/10.23887/ls.v25i1.18821

Sloman, K. N. (2010). Research trends in descriptive analysis. *The Behavior Analyst Today*, *11*(1), 20-35. https://doi.org/10.1037/h0100686

Surtees, V. (2016). Beliefs about Language Learning in Study Abroad: Advocating for a Language Ideology Approach. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *27*, 85-103. https://doi.org/10.36366/frontiers.v27i1.376

Turnbull, B. (2017). The translanguaging classroom: leveraging student bilingualism for learning. *International Journal of Bilingual Education and Bilingualism*, *22*(5), 639–640. https://doi.org/10.1080/13670050.2016.1275289

Vu, N. T., Chi, D. N., & Nguyen, H. T. M. (2025). Exploring Students' Intercultural Communicative Competence in the Light of English-Medium Instruction (EMI) in Higher Education in Vietnam. *Journal of Ethnic and Cultural Studies*, *12*(3), 152-175. https://doi.org/10.29333/ejecs/2168

Yen, P. H., & Danh, D. T. (2022). English–majored students' perceptions on teachers' use of Vietnamese in EFL classrooms. *JEELS (Journal of English Education and Linguistics Studies), 8*(2), 371-391. https://doi.org/10.30762/jeels.v8i2.3379

Yuvayapan, F. (2019). Translanguaging in EFL classrooms: Teachers' perceptions and practices. *Dil Ve Dilbilimi Çalışmaları Dergisi*, 15(2), 678-694. https://doi.org/10.17263/jlls.586811

Zumor, A. Q. A. (2019). CHALLENGES OF USING EMI IN TEACHING AND LEARNING OF UNIVERSITY SCIENTIFIC DISCIPLINES: STUDENT VOICE. *International Journal of Language Education*, 74-90. https://doi.org/10.26858/ijole.v1i1.7510

Bionote

Thinh Tien Vu is head of the Testing board and vice dean of the School of Languages, International University – VNU HCM. He received his M.A. in TESOL from the University of Canberra, Australia, and is currently a PhD candidate in English Language Studies and Teaching at NIDA, Thailand. He has 25 years of experience in teaching academic English, ESP, test preparation, and TESOL courses, and more than 12 years of experience in the roles of course/program designer and developer, academic trainer, and manager. His current research interests include EMI, Applied methodologies, Autonomy development, Assessment and Testing, Education management, and TELL.

Cam Diem Nhi Hoang is a senior undergraduate in English Linguistics and Literature at International University – VNU HCM. She currently teaches English to young learners and adults while actively engaging in academic inquiry. Her research interests involve EMI, teaching strategies, and the integration of 21st-century skills into language education.

#106 THE USE OF ARTIFICIAL INTELLIGENCE TOOLS IN ESSAY WRITING AMONG FIRST-YEAR ENGLISH MAJORS

Nguyen Phuong Thao[1]

Abstract

The use of Artificial Intelligence among students in their learning process has been attracting significant attention and investigation. The current study, accordingly, is an attempt to delve into the use of AI tools among first-year English majors to deeper understand the role of these tools in their essay writing process and the perceived effects on students' English writing skills. Data obtained from 60 questionnaires revealed that AI tools, ChatGPT and Google Gemini in particular, have become an integral part of the students in their essay writing across all stages, among which vocabulary, grammar, and ideas assistance are highly appreciated. In addition, the students tend to rely on AI tools in the writing process and generally believe on the positive short-term and long-term effects of AI tools on their English writing skills.

Keywords: Artificial Intelligence, first-year, English majors, essay writing

1. Introduction

1.1. Background of the study

The educational sector has been witnessing radical changes since the introduction of Artificial Intelligence (AI). The new technology has revolutionized the way students learn and provided them with brand new learning experiences through the use of AI tools. Almost in no time, these tools have become an integral part of students' learning process, making them students' new study companions.

1.2. Problem statement

Students majoring in English, accordingly, have utilized these tools in various academic tasks and purposes including word definition, pronunciation, examples along with translation and grammar-checking. In addition, in terms of academic courses, it is English writing that they seek support from these artificial companions the most as writing skills seem to be demanding to students generally. Furthermore, as observed by the researcher, the integration of AI tools in the student's writing process, especially in English essay writing, is becoming more and more common, which causes some concern regarding the writing competence of English majors.

1.3. Research objectives

Therefore, the current study is an attempt to explore first-year English-majors' practice of using AI tools in their English essay writing. Besides, their opinions concerning their levels of dependence on AI tools and their belief on the short-term and long-term impacts of these tools on their English writing ability can be revealed.

To achieve those aims, the following questions are to be addressed:

[1] *Faculty of Foreign Languages, HCMC University of Foreign Languages and Information Technology (HUFLIT), thaonp@huflit.edu.vn*

1. What are students' patterns of using AI tools in English essay writing?
2. What are students' opinions regarding their dependence on AI tools during their English essay writing process?
3. What are students' opinions regarding the short-term and long-term impacts of AI tools on their English writing ability?

2. Literature review

2.1. Students' patterns of using AI tools in English essay writing

2.1.1. The frequency of using AI tools in English essay writing

Research shows that AI tools are no longer occasional references but have become integral components of the English writing process for most university students in Vietnam, and English majors are no exceptions. In terms of essay writing, the reliance of students on these tools are universal, with AI being employed throughout the entire writing process on a frequent level (Bui & Tong, 2025; Lai & Trinh, 2025; Le, 2025). A study conducted at Can Tho university even found that 100% of English-majored participants utilize AI for their essay writing (Ho et al., 2024).

2.1.2. The AI tools used in English essay writing

Among the tools available, ChatGPT is the most dominant among students. Its popularity stems from its ability to simulate human-like conversations and provide instant feedback on a wide array of subjects**,** which facilitate students' idea generation, brainstorming, and content organization (Ho et al., 2024; Lai & Trinh, 2025; Pham, 2024; Tran, 2025). For more specific writing purposes, a range of tools are also employed including Quillbot (for paraphrasing), Grammarly (for grammar, spelling, and punctuation checks), Claude and Gemini (for content generation and searching for information) (Bui & Tong, 2025; Bui et al., 2022; Jen & Salam, 2024; Nguyen et al., 2025; Tran et al., 2025).

2.1.3. Stages of the essay writing process in which AI tools are employed

Moreover, AI tools are integrated across nearly all stages of the English essay-writing process, especially during the initial phrases where students need to establish a clear direction for their work. Accordingly, AI tools are most used for idea generation and brainstorming purposes, closely followed by defining topics, information seeking, outlining and structuring (Ho et al., 2024; Jen & Salam, 2024; Lai & Trinh, 2025; Le, 2025; Nguyen, 2025; Tran, 2025; Tran et al., 2025). In the drafting stage, AI provides students with sample paragraphs and essays, vocabulary enhancement, translation and paraphrasing (Bui et al., 2022; Bui & Tong, 2025; Lai & Trinh, 2025; Le, 2025; Ho et al., 2024; Nguyen et al., 2025; Tran et al., 2025). In the final stages, AI tools are used as automated evaluation systems to polish the work by checking grammar and mechanics, coherence and clarity and ensure it meets academic standards (Bui & Tong, 2025; Ho et al., 2024; Lai & Trinh, 2025; Le, 2025; Nguyen, 2025; Nguyen et al., 2025; Tran et al., 2025).

2.1.4. Students' treatment of AI-generated output

Students generally treat AI-generated essays as a supplementary resource or a writing companion rather than a final product. Therefore, they do not typically accept AI output blindly; instead, they critically assess the accuracy, relevance, and specificity of the information provided using different methods (Lai & Trinh, 2025; Nguyen et al., 2025; Tran et al., 2025). Besides, rather than directly copy-pasting content, the majority of students use AI

output as a trigger for brainstorming or a template to be modified by rephrasing AI-generated text to align it with their personal writing style and to avoid being flagged by plagiarism detection software (Tran et al., 2025). Importantly, most students emphasize that while AI can assist with structure and grammar, the development of core arguments and final organization remains their individual responsibility (Bui & Tong, 2025).

2.2. Students' opinions regarding their dependence on AI tools during their English essay writing process

University students demonstrate a complex relationship with AI, with a small percentage claiming high or full dependence while a much larger majority expresses a significant fear of becoming over-reliant.

According to study at a Hanoi university, 9,6% of the English majors admitted always using ChatGPT for academic writing while 4,3% of them heavily relied on AI tools to complete all their writing assignments (Lai & Trinh, 2025). Significantly, 45% of interviewed students at Thu Dau Mot University acknowledged a dependence on AI-generated answers, sometimes failing to critically evaluate the output (Nguyen et al., 2025). Also in this study, a significant positive correlation between the frequency of AI usage and the recognition of over-reliance was found.

The other students, who do not consider themselves currently dependent, often self-identify as being at high risk for it. A study by Bui and Tong (2025) pointed out that the single highest risk associated with AI tools was dependence on technology, cited by 79.8% of students. In other studies, the fears of overreliance on AI and its possible consequences including lower independent writing ability, diminished problem-solving skills and critical thinking erosion are prevalent among students (Bui & Tong, 2025; Lai & Trinh, 2025; Ho et al., 2024; Tran et al., 2025). To prevent themselves from the possible risks, students struggle to treat AI as a companion or guide as a specific strategy (Le, 2025; Tran et al., 2025).

2.3. Students' opinions regarding the short-term and long-term impacts of AI tools on their English writing ability

While being consistent regarding the benefits given by AI tools in the short term, students are deeply divided on how these tools impact their long-term writing skills. For the immediate, short-run benefits, it is widely believed that AI tools generally enhance their writing ability by reducing anxiety, boosting motivation, giving useful feedback, enhancing their English abilities, assisting them in better paragraph and essay planning and keeping track of their own progress (Le, 2025; Nguyen, 2025). Students, however, express uncertainty regarding long-term retention of the new academic vocabulary or grammatical mastery they gained through AI assistance and long-term dependency, which may result in cognitive and skill erosion (Bui & Tong, 2025; Lai & Trinh, 2025; Tran et al., 2025).

3. Methodology

3.1. Participants

The population of this study comprises of 60 HUFLIT first-year English majors who were taking English Essay Writing course, which the researcher was in charge of.

3.2. Research instruments

Questionnaire is the major research instrument employed in this study. The questionnaire is written in English and is divided into three main parts exploring different aspects regarding first-year English majors' practice of using AI tools in English essay writing

and their opinions concerning dependence level and the short-term and long-term impacts of these tools on their English writing ability. Different question types are utilized across the question items to provide the researcher with varying levels of details, accuracy, and meaningful insights into the investigated issues.

In addition, classroom observation is included. The researcher during teaching sessions recorded students' practice of using AI tools in their essay writing process to recognize their usage patterns and level of dependence on these tools.

3.3. Data collection and analysis

The data was collected within two weeks, after which SPSS was employed in the analysis process. The findings, accordingly, were presented in the forms of tables.

4. Findings

4.1. From questionnaires

4.1.1. Students' patterns of using AI tools in English essay writing

Table 1

The frequency of using AI tools in essay writing

AI usage	Frequency	Percent
Always	6	10%
Often	34	56.7%
Rarely	1	1.7%
Sometimes	19	31.7%

Table 2.

AI tools used in essay writing

AI Tools	Percent
ChatGPT	75%
Grammarly	20%
QuillBot	3.3%
Google Gemini	66.7%
Other	11.7%

As shown in Table 1, AI tools are frequently incorporated in the students' English essay writing process as more than two thirds of the investigated English majors use them on a regular basis. Among the tools available, ChatGPT is the dominant, which is closely followed by Google Gemini (Table 2).

Table 3.

Essay writing stages with AI support

Writing Stages	Percent

Brainstorm	55%
Create outlines	48.3%
Write first draft	26.7%
Check grammar and vocabulary	81.7%
Revise and edit	51.7%
Cite	13.3%

Table 4.

Students' essay writing process

Essay writing process	Frequency	Percent
I rarely or never use AI in my writing process.	1	1.7%
I write first, then use AI to improve or correct my writing.	26	43.3%
I use AI to generate ideas, then write the essay mostly by myself.	7	11.7%
I use AI to create an outline, then develop it on my own.	17	28.3%
I use AI to generate a full draft, then revise and edit it.	3	5%
I use AI throughout the entire writing process.	5	8.3%
Other	1	1.7%

The students tend to use AI tools mainly for checking grammar and vocabulary. The other purposes include brainstorming ideas, revising and editing, and creating outlines (Table 3). In addition, as revealed by Table 4, the most common procedure in writing an English essay, as applied by almost half of the participants, is making their own essay first, after which AI tools are used for polishing purpose. The second common procedure, applied by 28.3% of the students, is using AI for outline developing, after which the essay is written by the students themselves.

Table 5.

Revision frequency after using AI tools

	Minimum	Maximum	Mean	Std.Deviation
Revision frequency	2	5	3.9	0.72952

Table 6.

Revised aspects after using AI tools

Revised aspects	Percent
I correct grammar and vocabulary based on AI feedback	70%
I rewrite sentences to make them sound more natural	55%

I reorganize ideas or paragraph structure	36.7%
I simplify or clarify ideas	23.3%
I combine my own ideas with AI-generated content	41.7%
I remove parts I think are incorrect or unnatural	20%
I make only minor changes	1.7%
Other	3.3%

Tables 5 shows that students' revision of their own essay after seeking AI assistance is a frequent practice (M=3.9, SD=0.72952). They do this in different ways, among which correcting grammar and vocabulary mistakes is the most common, closely followed by rewriting sentences and reorganizing the ideas and paragraphs (Table 6).

4.1.2 Students' opinions regarding their dependence on AI tools during their English essay writing process

Table 7.

Students' dependence on AI tools during essay writing process

	Minimum	Maximum	Mean	Std.Deviation
AI Dependence	2	5	3.45	0.832

Table 8.

Students' claim of their final essay

Claim of Authority	**Frequency**	**Percent**
Completely my own work	11	18.3%
Mostly my work with some AI assistance	22	36.7%
A combination of my work and AI	24	40%
Mostly AI-generated	1	1.7%
Completely AI-generated	2	3.3%

As presented in Table 7, on average, students tended to agree that they depend on AI tools during the English essay writing process (M=3.45. SD=0.832). However, with more than three fourths of the participants claiming that the final essay is either mostly their work with AI assistance or a combination of their work and AI, the first-year English majors do not generally believe AI is the primary author of their essays (Table 8).

Table 9.

The relationship between students' AI usage frequency and dependence level

			AI Usage Frequency	**Dependence**

Spearman's rho	AI Usage Frequency	Correlation Coefficient	1.000	.367
		Sig. (2-tailed)		.004
	Dependence	Correlation Coefficient	.367	1.000
		Sig. (2-tailed)	0.004	

Nevertheless, the correlation analysis shown in Table 9 suggests that more frequent AI use is associated with greater perceived dependence on AI. Therefore, while students generally maintain a sense of authorship, increased AI use may lead to increased reliance on AI tools during the writing process.

4.1.3. Students' opinions regarding the short-term and long-term impacts of AI tools on their English writing ability

Table 10.

Students' claim of their writing skills improvement after using AI tools

	Minimum	Maximum	Mean	Std.Deviation
Writing skills improvement	2	5	3.7667	0.72174

Table 11.

Improved aspects after using AI tools

Learn from AI Tools	**Percent**
New vocabulary	70%
Grammar structures	70%
Essay organization	46.7%
Idea development	71.7%
Nothing significant	5%
Other	1.7%

As indicated in Table 10, students generally perceived that AI tools contributed positively to the improvement of their English writing skills in the short term (M=3.7667, SD=0.72174). The most improved aspects include idea development, vocabulary, grammar, and essay organization (Table 11).

Regarding the long-term effects of AI tools, data from the open-ended question unfolded similar attitudes from the participants as they tend to believe that AI tools are good for their English writing skills along the way. However, many of them emphasized that these benefits come from the appropriate use of these tools only since misuse or over-dependence may bring about significantly harmful impacts.

4.2. From observation

During English essay writing sessions in the classroom, it was recorded that AI tools, ChatGPT and Google Gemini in particular, are frequently used by the first-year English majors

of the two classes the researcher was in charge of. Even though the students were encouraged to brainstorm ideas themselves, many were stuck in the process and had to seek for AI assistance. In addition, AI tools like ChatGPT are acted as their encyclopedia where they inquire all aspects related to the essay including the topic-related knowledge, vocabulary and grammar, etc.

More importantly, the level of dependence on AI tools varies greatly according to the English proficiency of students. Those who are good at English, especially at English writing skills, rarely ask for AI support. They tend to use these tools only at the final stage of the process when they want their essay to be polished. The relatively weak students, meanwhile, heavily depend on AI tools, otherwise their essays would be full of mistakes and irrelevant ideas.

5. Discussion

5.1. Students' patterns of using AI tools in English essay writing

There are similarities between the students in the current study with the students in previous research regarding the practice of using AI tools in English essay (Bui & Tong, 2025; Lai & Trinh, 2025; Le, 2025). They generally use a range of AI tools on a regular basis during English essay writing process for different purposes and do revise their own essay after receiving AI feedback.

However, there are quite several differences. Firstly, although ChatGPT continues to be the first choice of the students, Gemini is gaining its popularity considerably as the percentages of the students using ChatGPT and those using Gemini are quite close. Secondly, while the students in other studies tend to utilize AI in their very first stage of essay writing, which is ideas generation and brainstorming, the Huflit first-year English majors prioritize vocabulary and grammar aspects (Ho et al., 2024; Jen & Salam, 2024; Lai & Trinh, 2025; Le, 2025; Nguyen, 2025; Tran, 2025; Tran et al., 2025). And it is also vocabulary and grammar mistakes that they tend to revise firstly after seeking AI feedback.

Besides, this study points out the procedure the students apply in their English essay writing, which has been under little investigation among previous research. The most common procedure is where students make their own essays first, after which they seek for AI polish. Another alternative is where AI was the one to generate the outlines from which the students develop the whole essays themselves.

5.2. Students' opinions regarding their dependence on AI tools during their English essay writing process

The students' self-report of their dependance on AI tools during essay writing process pointed out a comparatively clear-cut answer compared to what research has consistently pointed out. The first-year English majors tend to admit their dependence on AI tools during the English essay writing process, yet interestingly, most students do not regard AI as the primary author of their essays. However, the current study confirmed the positive correlation between the frequency of using AI tools and the levels of dependence, which Nguyen et al.(2025) previously figured out.

5.3. Students' opinions regarding the short-term and long-term impacts of AI tools on their English writing ability

On the one hand, the participants' views on the short-term effects are in line with findings from relevant research. They generally appreciate the AI assistance which lead to their

English writing improvement, especially on ideas, vocabulary and grammar aspects (Bui & Tong, 2025; Nguyen, 2025; Nguyen et al., 2025; Pham, 2024).

On the other hand, while the students in previous research express uncertainty regarding long-term retention of the new academic vocabulary or grammatical mastery they gained through AI assistance and the general writing improvement, the participants in this study are quite positive about such effects. However, like students in those research, the prospects of long-term dependency, which may result in cognitive and skill erosion, still concern them to a certain extent (Bui & Tong, 2025; Lai & Trinh, 2025; Tran et al., 2025). Hence, though showing positive attitude about the long-term impacts, many of the students insist that in order to bring about these benefits, the right use of AI tools must be guaranteed, otherwise the bad consequences are likely to happen.

6. Conclusion

AI tools, especially ChatGPT and Gemini, have become integral components of the English essay writing process among first-year English majors. They use these tools on a regular basis yet treat them as a supplementary resource or a writing companion rather than a final product to be submitted without modification. In English essay writing procedure, AI is often acted either as polishing tools after students have finished their own essays or as an outline generator from what students develop their own writing.

Grammar and vocabulary are the two main language areas that are of great importance to first-year English majors as they are repeatedly report to be not only students' main purposes for seeking AI assistance but also the language areas they get improved the most in the short term along with idea development, and accordingly making necessary changes in their final essays. Regarding the long-term effects of AI tools on their English writing ability, students remain a positive attitude, yet many insist that these benefits only present if students use AI tools in the right way, otherwise critical thinking skills and autonomy will be negatively affected.

Though tending to admit their dependence on AI tools during the process of English essay writing, the students generally claim their authorship of the final essay. However, given the fact that a positive correlation between the frequency of using AI tools and dependence was figured out, and students' reliance on these tools, especially those with low English proficiency, is confirmed by classroom observation, strategies to help students use AI tools effectively without becoming over-dependent is of utmost importance. First and foremost, teachers need to make sure that there is an unplugged writing session when students need to brainstorm and make the first draft on their own. Pair work, groupwork and whole class teaching may find themselves helpful at this stage as these collaborative activities support all students in the writing process, especially those with low English proficiency. AI assistance may follow afterwards for better ideas, organization, vocabulary or grammar, yet the essence of the essay, which is the students' original outline, should be preserved. Furthermore, students should take note of what they have learned from AI to ensure both short-term and long-term writing improvement.

Despite having gained insight into the use of AI tools in the process of English essay writing among first-year English majors, this study has a quite small population and a limited number of research instruments used. Therefore, future research should include more participants and incorporate a variety of tools to enhance validity and generalization. Besides, future research could investigate more about AI usage in English writing among students of

different English levels, from which strategies to maximize AI tools' positive effects on students can be suggested.

References

Bui, T. U., & Tong, V. A. (2025). The impact of AI writing tools on academic integrity: Unveiling English-majored students' perceptions and practical solutions. *AsiaCALL Online Journal*, *16*(1), 83-110.

Bui, V. H., Hoang, M. H., & Cu, T. H. (2022). Employing the QuillBot application in order to sharpen paraphrasing skills in writing academic essays for English-majored students at the School of Foreign Languages - Thai Nguyen University. *TNU Journal of Science and Technology*, *227*(13), 116–124.

Ho, B. N., Ho, B. N., & Thai, C. D. (2024). EFL students' perceptions and practices of using ChatGPT for developing English argumentative essay writing skills. *European Journal of Alternative Education Studies*, *9*(1), 168–216.

Jen, S. L., & Salam, A. R. (2024). Using artificial intelligence for essay writing. *Arab World English Journal*, 90–99.

Lai, K. L., & Trinh, H. L. (2025). The effects of ChatGPT on academic writing proficiency among English-majored undergraduates at a university. *VNU Journal of Foreign Studies*, *41*(3), 92–109.

LÊ, H. (2025). ENHANCING LEARNER AUTONOMY IN ENGLISH WRITING SKILLS THROUGH THE ADOPTION OF CHATGPT: A CASE STUDY OF ENGLISH-MAJORED FRESHMEN AT HO CHI MINH UNIVERSITY OF FOREIGN LANGUAGES AND INFORMATION TECHNOLOGY (HUFLIT). TESOL International Conference 2025.

Hoang, N. H. (2025). Chatgpt-assisted language learning: effects on vietnamese english majors' writing skills and motivation. The JALT CALL Journal, 21(2), 102429. https://doi.org/10.29140/jaltcall.v21n2.102429

Thi, X. H. N., Thien, H. V. H., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through ai-powered tools: perceived benefits and challenges among vietnamese EFL students. Discover Education, 4(1). https://doi.org/10.1007/s44217-025-00905-9

Pham, Q. T. (2024). Chatgpt's influence on english-majored students' writing: insights from lecturers at a university in hanoi. Tạp chí Thiết bị Giáo dục.

Tran, H. N., Le, T. N., & Tran, B. U. (2025). AI tools in learning academic writing: Benefits and challenges for MA students in the English language studies at the Industrial University of Ho Chi Minh City. *International Journal of AI in Language Education*, *2*(1), 74–91.

Tran, N. M. (2025). Factors affecting the use of ChatGPT in academic writing perceived by English-majored students. *VNU Journal of Foreign Studies*, *41*(3), 110–125.

#107 THE INTRINSIC MOTIVATION OF GENERATION Z AT COFER IN VIETNAM'S SHIFTING ESL CONTEXT

Le Nguyen Hoang Nam[1]

Abstract

In 2023, Vietnam formally redesignated English as a Second Language, raising important questions about what drives Generation Z learners in this new context. This study examined the intrinsic motivation of 83 students at the College of Foreign Economic Relations (COFER), where many engaged extensively with English independently, yet hesitated to speak in class. This study applied thematic analysis of structured written interviews, based on established theories such as Self-Determination Theory (Deci & Ryan, 1985), Dörnyei's L2 Motivational Self System (2009), Norton's Identity Investment Framework (2000), and the Willingness to Communicate model (MacIntyre, Clément, Dörnyei, & Noels, 1998). Using these frameworks, the analysis identified five connected themes. These are enjoyment-driven engagement, identity investment, digital media as an intrinsic motivational space, autonomous learning orientation, and the motivational-communicative gap. The findings indicate that students were considerably more intrinsically motivated than previously expected, with English closely linked to their ambitions, identities, and digital experiences. However, foreign language anxiety remained a significant barrier, and no direct link was found between high motivation and willingness to communicate face-to-face. The paper proposes implications for pedagogical practice, AI integration in learning environments, and student assessment within Vietnam's evolving ESL context.

***Keywords**:* Self-Determination Theory; L2 Motivational Self System; digital informal learning; willingness to communicate; foreign language anxiety; thematic analysis

1. Introduction

Following a 2023 policy change in Vietnam, English has been formally redesignated as a Second Language (ESL), moving away from its prior classification as a Foreign Language (EFL). The significant impact of motivation on second language acquisition is widely recognized (Gardner, 1985; Dörnyei, 2009). Traditionally, pragmatic and extrinsic elements have been dominant influences in Vietnamese educational contexts (Nguyen & Habók, 2021; Le & Nguyen, 2023).

However, this perspective is becoming less reflective of Generation Z learners, who were born between 1997 and 2012 (Seemiller & Grace, 2016). Gen Z learners frequently engage with English content on platforms such as TikTok and YouTube without institutional encouragement (Sockett, 2014). Despite a growing interest in Vietnamese language studies, the intrinsic motivation of this generation requires further investigation. Existing research has relied predominantly on large-scale quantitative surveys of general undergraduate or English-

[1] *Faculty of Foreign Languages, College of Foreign Economic Relations (COFER), Ho Chi Minh City, Vietnam, Email: lnhnam@cofer.edu.vn*

major populations (Nguyen & Habók, 2021; Thu & Lee, 2024). Therefore, the intrinsic motivations of 83 Gen Z students enrolled at the College of Foreign Economic Relations (COFER) were explored to address this gap. Thematic analysis was conducted with guidance from Self-Determination Theory (Deci & Ryan, 1985), the L2 Motivational Self System (Dörnyei, 2009), Identity Investment (Norton, 2000), and Willingness to Communicate (MacIntyre, Clément, Dörnyei, & Noels, 1998). Building upon these theoretical frameworks, this study seeks to address the following research questions:

RQ1. What intrinsic motivational themes - including enjoyment, identity investment, and digital media engagement - characterize Gen Z students' English learning experiences at COFER?

RQ2. How do these students perceive Vietnam's shift in ESL policy, and how does it interact with their intrinsic motivation?

2. Literature Review

2.1. Intrinsic Motivation and Self-Determination Theory

This study adopts Self-Determination Theory (SDT; Deci & Ryan, 1985, 2000) as its central theoretical framework. SDT proposes a motivational spectrum extending from amotivation through varied forms of extrinsic regulation to full intrinsic motivation (Ryan & Deci, 2000). Internalization is understood as the procedure wherein externally controlled actions are progressively converted into self-governed behaviors (Deci & Ryan, 2000). Within existing scholarly works (Le & Nguyen, 2023; Nguyen, 2022; Thu & Lee, 2024), Vietnamese learners' motivation has been examined predominantly through survey instruments administered to general undergraduate or English-major cohorts, with career-related instrumentality frequently foregrounded. These inferences have generally been derived from quantitative measures. Rahaman (2025) argues that this positivist orientation has left the social and structural dimensions of language learning underexamined, and advances investment as a sociological counterpart to the psychological construct of motivation. Nevertheless, the multifaceted motivational contexts of digitally active younger learners are often not comprehensively captured by such methodologies. An early Vietnamese exception is Tran (2007), whose qualitative study of thirty English majors found intrinsic concerns — interest, passion and inspiration — to be more salient than the institutional and linguistic needs emphasised in the surrounding literature. Accordingly, this study adopts a qualitative methodology to illuminate dimensions that survey-based research has tended to overlook.

2.2. Dörnyei's L2 Motivational Self System

Dörnyei's (2009) L2 Motivational Self System (L2MSS) reconceptualizes language acquisition as an identity-driven process. The L2MSS consists of the Ideal L2 Self, the Ought-to L2 Self, and the L2 Learning Experience. The Ideal L2 Self holds particular motivational significance: a clear and emotionally resonant ideal self-image has been shown to predict learning effort more powerfully than conventional instrumental motivation (Dörnyei, 2009). This result is supported by Dörnyei and Ushioda (2011). Digital environments are recognized to have a considerable effect on the evolution of this aspirational self-image for Gen Z learners. Within the Vietnamese setting, the key motivational effects of Ideal L2 Self orientations, which are associated with desired career goals and international interaction, were identified by Thu and Lee (2024), even amidst significant societal expectations. Despite this, identity-based motivation has not been sufficiently explored within Vietnamese EFL research, a gap that this study seeks to address.

2.3. Identity Investment in Digital Spaces

Norton's identity investment theory (2000; 2013) has prompted a reconceptualization of language learning as a socially constructed endeavor of self-formation. Within this framework, learner investments are understood to extend beyond linguistic codes to encompass their aspired future selves. Investments are made in the pursuit of an ideal social identity, integration into communities of practice, and projections into imagined futures.

Subsequently, this framework has been broadened into digital realms by Darvin and Norton (2015). Streaming services and social media platforms are now regarded as primary sites for identity investment among contemporary learners. For Vietnamese Gen Z students, English in these digital environments is not perceived as a mere school subject but as a vehicle for self-expression. Multifaceted identities are constructed online, particularly by individuals who may experience a lack of recognition within traditional classroom settings.

Evidence from Vietnamese higher education lends support to this phenomenon. Thu and Lee (2024) found the Ideal L2 Self to be the strongest predictor of intended learning effort among Vietnamese undergraduates, outweighing externally imposed obligation. Working within Darvin and Norton's expanded model, Rahaman (2025) further demonstrates how imagined identity and group membership shape commitment to English in an Asian EFL setting marked by uneven access to resources. This convergence of identity and autonomy is directly witnessed in COFER's Gen Z learners.

2.4. Willingness to Communicate and Foreign Language Anxiety

The WTC model proposed by MacIntyre et al. (1998) views communication as an active, shifting process influenced by a mix of personal confidence, intrinsic motivation, and broader attitudes. Crucially, this framework points out that the link between wanting to learn and actually speaking is rarely perfect. A learner can be deeply motivated on an intrinsic level, yet still lack the readiness to initiate a conversation. Foreign language anxiety, first characterised by Horwitz, Horwitz and Cope (1986) as a distinct situation-specific anxiety comprising communication apprehension, test anxiety and fear of negative evaluation, can suppress communication in ways that mirror the behavioural restrictions imposed by classroom-based anxiety (Peng, 2014). Cao and Philp (2006) similarly found that willingness to communicate varied with the interactional context of the classroom: learners participated less readily in whole-class settings than in small groups or pairs, and identified group size, familiarity with interlocutors and self-confidence among the factors shaping that decision. This pattern is directly relevant to the findings of the present study.

2.5. Digital Informal Learning and Gen Z Learners

Digital environments host dynamic learning ecosystems that exist beyond institutional control. Sockett's 2014 work on Online Informal English Learning (OILE) conceptualized this, with streaming platforms and music communities being examples. Out-of-class technology use of this kind is documented in Lai's (2017) study of Hong Kong learners, which shows such use to be sustained by learners themselves and shaped by an interaction of psychological and sociocultural factors rather than by institutional direction. Contributors to Dressman and Sadler's (2019) handbook further emphasise that informal digital spaces especially promote self-governed participation, attributable primarily to the lack of institutional restrictions that often impede self-motivated learning.

This trend is particularly evident among Gen Z, for whom digital literacy is regarded as an inherent attribute rather than a learned skill (Seemiller & Grace, 2016). The same authors

report a propensity for investigative, self-directed learning that bypasses traditional pedagogical methods. In Vietnam, these generational traits converge with a vibrant digital setting where English has become an integrated element of daily cultural exchange, rather than simply a separate academic subject.

3. Methodology

3.1. Research Design

Within Vietnam's developing ESL context, the inherent drive of Gen Z college students to learn English was examined through qualitative research. Due to difficulties in quantifying subjective interpretations and personal narratives, a qualitative method was deemed suitable for investigating motivational experiences during English language acquisition (Creswell & Poth, 2018). This study employs thematic analysis, conducted utilizing the six-phase framework of Braun and Clarke (2006). Meaning patterns within the qualitative data are systematically identified, analyzed, and interpreted through this framework. Therefore, a richer understanding of the motivational factors influencing Gen Z students was sought.

3.2. Research Context

The research was conducted at the College of Foreign Economic Relations (COFER) in Ho Chi Minh City. As a vocational institution, COFER offers expertise in economics, foreign languages, and international trade, drawing students from across Southern Vietnam, primarily from Ho Chi Minh City and adjacent provinces. This study chose COFER as its primary setting for two main reasons. First, the institution treats English not merely as a standard language requirement, but as a crucial condition for graduation and future career success. Second, its student body represents the typical urban Gen Z in Vietnam: they are tech-savvy, highly driven, and actively balancing local traditions with global trends.

3.3. Participants

Eighty-three Gen Z students from COFER, participating in the 2025–2026 academic year, were included in the sample. A variety of disciplines were represented, with Logistics (n=33) and E-commerce (n=26) being the most prevalent. Business Administration, Finance and Accounting, and Import/Export were represented by smaller numbers (n=11, n=11, and n=2, respectively). The second-year cohort constituted the largest group (n=60), followed by third-year (n=13) and first-year students (n=10). Limited English proficiency was commonly reported among participants. Fifty participants reported a basic proficiency level, 29 reported intermediate, and only four indicated upper-intermediate or higher. These results are considered representative of the language skills generally observed in Vietnamese university students outside specialized language programs.

3.4. Data Collection Instrument

Data were collected through a structured written interview comprising 16 open-ended questions administered in Vietnamese to maximize participants' expressive capacity and minimize linguistic barriers to self-disclosure. Questions were sequenced to elicit rich, reflective narrative responses across five motivational dimensions: enjoyment and intrinsic engagement, identity investment, digital media engagement, autonomous learning orientation, and perceptions of Vietnam's ESL policy shift. Representative questions included: "Tell me about a time you used or learned English simply because you enjoyed it, not because of any pressure. How did you feel?"; "If there were no grades or requirements, would you still learn English on your own? Why?"; and "Use ONE sentence to describe your relationship with

English." The structured written interview was administered online via Google Forms between 26 March and 2 April 2026, and participation was entirely voluntary.

3.5. Data Analysis

The study adopted Braun and Clarke's thematic analysis. Working with 83 Vietnamese responses required a careful approach; therefore, the initial coding for both surface and latent meanings took place entirely in the original language. English translations were reserved solely for essential quotes during the final write-up.

The coding process yielded five distinct themes. On one hand, the data highlighted enjoyment-based engagement, identity investment, and the tendency to treat digital media as an intrinsic motivational environment. On the other hand, it revealed a stark disconnect between feeling motivated and actual communication, despite a strong orientation toward autonomous learning. To contextualize these specific behaviors, the results were mapped against Norton's identity framework, SDT's motivational continuum, Dörnyei's Ideal L2 Self, and the WTC model.

3.6. Positionality and Trustworthiness

The researcher's role as a lecturer at COFER, with direct pedagogical experience with the participant population, affords interpretive advantages; however, this positioning also raises potential concerns regarding reflexivity. To enhance the trustworthiness of the analysis, several measures were undertaken. Member checking was approximated through informal follow-up discussions with a subset of participants to verify the accuracy of theme interpretations. Thick description (Geertz, 1973) was employed throughout the Findings section to ensure that the analytical claims are grounded in participants' own words rather than imposed interpretive frameworks. Additionally, the researcher maintained a reflexive journal throughout the analysis process to monitor and document potential biases arising from his institutional positioning.

4. Results

Thematic analysis of structured written interviews from 83 Gen Z students at COFER revealed five intrinsic motivational themes. Although extrinsic pressures were present, intrinsic orientations emerged as the dominant motivational force. Presented below are the themes and example quotations from participants, translated from Vietnamese.

Theme 1: Enjoyment-Driven Engagement

Among the collected data, a prominent theme was the role of enjoyment as a key motivator in English language learning. Experiences with English were often driven by intrinsic engagement, detached from academic needs or external pressures. This aligns with Deci and Ryan's concept of pure intrinsic motivation, as evidenced by a significant number of participants.

Participants consistently described positive affective states, joy, excitement, pride, and comfort when recounting experiences of voluntary English use. One participant captured this quality of effortless engagement particularly vividly:

> *"I felt happy and relaxed, as if I were discovering something new rather than 'studying'. […] After that, I realised that when I learn because I like it, I take things in faster and remember them far longer."*

One participant observed that the exhilaration of gameplay overshadowed the acquisition of 40 vocabulary words in a single hour. This experience stood in stark contrast to conventional study methods. Joy was also expressed by individuals upon comprehending foreign film dialogue without subtitles, resulting in substantial pride and satisfaction. A powerful drive to continue learning was cultivated by such instances of sudden comprehension, where formerly difficult content became understandable independently. Significant advancements in understanding yielded gratification.

Notably, enjoyment in this theme was not passive but generative. Participants described a self-sustaining motivational cycle in which pleasurable experiences with English prompted further voluntary engagement. As one second-year student reflected:

> *"I realized that when I learn because I enjoy it, everything becomes lighter and more interesting. There's no sense of being forced just excitement and motivation from within. That was the first time I truly felt the joy of learning for myself."*

Theme 2: Identity Investment

The second theme emphasizes the considerable personal investment made by participants in English. It is viewed as essential for their desired professional and social identities. English was portrayed by many not as a neutral tool, but as integral to their envisioned selves. This aligns with Norton's identity investment concept (2000) and Dörnyei's Ideal L2 Self construct (2009), reflecting the deep motivational significance of identity-based frameworks in this context.

Participants frequently articulated aspirational self-images in which English proficiency was inseparable from their envisioned futures:

> *"English is not just communication for me it is a bridge that connects me to international friends and opens up opportunities I couldn't access otherwise."*

> *"I want to apply to foreign companies because I believe the environment shapes your character and working style. That's why I actively learn English not just to get a job, but to become the kind of person I want to be."*

The investment in English was regarded by many participants as closely tied to aspirations of global inclusivity and cultural engagement, along with a desire to connect with communities beyond Vietnam. It was noted by one participant that, "While exploring TikTok and witnessing people from diverse countries communicating seamlessly in English, a longing to be part of that world arises within me. This drive significantly enhances my motivation for learning, surpassing any academic accomplishments."

Notably, identity investment in this theme did not yield uniformly positive outcomes. Several participants reported experiencing anxiety related to their identities due to perceived discrepancies between their current abilities and their envisioned Ideal L2 Self. One metaphor effectively illustrated this tension: "English and I are familiar, yet we have not developed a mutual affinity."

Theme 3: Digital Media as Intrinsic Motivational Space

Within the COFER framework, digital platforms like TikTok and YouTube are considered pivotal sites where intrinsically motivated English engagement is observed among Generation Z learners. It was reported by a significant number of participants that English-language digital content is voluntarily sought after outside of formal instruction. This behavior

is motivated by curiosity, enjoyment, and a desire for communities and content inaccessible in Vietnamese.

Digital platforms were characterized by participants not merely as educational tools, but rather as environments fostering genuine and enjoyable interactions where the acquisition of English emerged organically from these experiences. One individual remarked, “English content on TikTok and YouTube captivates me not for study purposes, but for enjoyment. Over time, I became aware of my natural progress in pronunciation and vocabulary.” Other participants noted, “Each instance of comprehending a joke in an English video without translation fills me with joy. This sense of achievement motivates my continued engagement with the content.”

Social media platforms were recognized as important motivational environments that facilitated the observation of genuine English usage by native speakers. One participant noted, "The sight of individuals my age, or even younger, speaking English fluently as if it were their native language inspires both admiration and a desire to achieve the same proficiency." This highlights the powerful influence of authentic interactions on language learning aspirations.

Within this theme, the importance of AI tools, particularly ChatGPT, was underscored. It was indicated by numerous participants that AI was being utilized as a conversational companion for voluntary English practice, instead of for academic purposes. Conversations with ChatGPT afforded participants the ability to make errors without embarrassment, resembling the support of a patient and non-judgmental practice partner.

Theme 4: Autonomous Learning Orientation

Participants demonstrated a robust inclination toward self-governed English acquisition, irrespective of imposed requirements or external encouragement. The vast majority of participants affirmed that they would continue learning English even without grades or formal requirements. Many articulated a nuanced understanding of English as a pursuit of personal significance, rather than merely a duty imposed by institutions.

Participants described a wide range of autonomous learning strategies from passive exposure through music and film to active self-study through vocabulary apps, AI-assisted writing practice, and self-initiated speaking exercises:

> *"If there were no grades, I would still learn English. For me, it's not just a subject it's a tool that helps me connect with the world and understand things I couldn't access otherwise."*

> *"I roleplay in my head, imagining I'm traveling abroad and speaking English on my own. It's my way of practicing without anyone knowing."*

Several participants demonstrated what SDT theorists would recognize as integrated regulation the highest form of extrinsic motivation internalization, in which English learning has become so personally meaningful that it is experienced as an expression of the self rather than an external requirement:

> *"Learning English is part of who I am now. I don't think of it as studying anymore, it's just something I do because I want to grow."*

Theme 5: The Motivational-Communicative Gap

The fifth and theoretically vital theme involved a paradox in participant motivation. High intrinsic motivation for English learning was identified. However, a significant reluctance

to participate in genuine, face-to-face English communication was consistently indicated by participants. This motivational-communicative discrepancy, where learning motivation did not easily become a willingness to communicate, became one of the most widespread and emotionally significant trends throughout the data.

Participants described a range of anxiety-inducing situations in which their intrinsic motivation was effectively neutralized by fear of negative evaluation, embarrassment at pronunciation errors, and perceived inadequacy in real-time communication:

"I know English, and English knows me but we don't get along."

"We are two parallel lines that can never meet."

"I have everything but not English."

More concretely, participants described specific situations in which anxiety overrode motivation:

"When I'm with foreign customers at work, I suddenly can't remember anything. I know what I want to say in my head, but I can't get the words out. I feel so ashamed."

"In class, I prepared what I wanted to say but when the teacher called on me, I froze. Everyone was looking at me and I just went blank."

Significantly, the motivational-communicative gap was not experienced as a motivational endpoint but as a source of productive tension for many participants. Several described shame and communicative failure as catalysts for renewed intrinsic engagement:

"After that embarrassing moment, I went home and practiced for hours not because anyone told me to, but because I refused to feel that way again."

Looking at these results through MacIntyre et al.'s WTC framework reveals a complex emotional layer for Vietnamese Gen Z learners. For this demographic, the jump from feeling motivated to actually communicating is heavily influenced by their emotions. Interestingly, foreign language anxiety does not act strictly as a roadblock here. In certain situations, this anxiety actually drives them to work harder and sustains their intrinsic motivation to learn.

Perceptions of Vietnam's ESL Policy Shift

Participants were also asked directly what they made of English becoming a Second Language in Vietnam, and whether this changed their motivation. Almost all responded. Their accounts converged on a reading of the policy as raising the stakes attached to English rather than as changing how English is taught. Many described the redesignation as an inevitable consequence of international integration, and a recurring formulation was that English would cease to be an advantage and would instead become a requirement.

Many said the shift would raise their motivation, though the increase was not uniformly comfortable. Several named pressure directly, and a recurring concern was the fear of being left behind. For some participants, the policy changed the character of their motivation rather than its intensity:

"It shifts my motivation from studying to pass exams towards studying in order to live and work, because English then becomes a compulsory survival skill rather than a choice."

"Those who are not yet good at it may feel they are being 'left behind'."

In answering this item, participants spoke about opportunity, requirement and social standing. They did not raise classroom routines, assessment formats or teaching practice.

5. Discussion

The findings of this study reveal a motivational landscape among Gen Z students at COFER that is richer, more complex, and more intrinsically oriented than existing Vietnamese EFL research would suggest. This section discusses the five identified themes in relation to the theoretical frameworks outlined in the Literature Review, draws connections across themes, and addresses the study's two research questions.

5.1. Intrinsic Motivation as the Dominant Orientation

Within the motivational narratives of the participants, the consistent dominance of intrinsic motivation has been identified as the most theoretically significant finding. This finding directly challenges prevailing characterizations of Vietnamese learners. With the vast majority of participants indicating willingness to learn English independently of academic requirements, the data suggest that Gen Z students at COFER have already developed what SDT theorists would recognize as identified or integrated regulation forms of self-determined motivation in which English learning is experienced as personally meaningful and congruent with the self rather than externally imposed (Deci & Ryan, 2000).

Despite the impact of external elements, the original finding retains its significance. Extrinsic and intrinsic motivations are understood to function simultaneously across regulatory levels, as recognized by SDT's motivational continuum. For numerous individuals, external pressures were seen to serve as an interim mechanism of introjected regulation, offering temporary support. Intrinsic motivation was observed to coexist with, and at times enhance, deeper motivational internalization rather than being overshadowed by external pressures. This interplay is particularly highlighted in Theme 5, where efforts in self-directed learning were often revitalized by shame and communication breakdowns instead of being diminished.

5.2. Identity as a Motivational Engine

Identity investment's motivational role emerges through aspirational self-images, desires for global connection, and narratives of English as a vehicle for self-expression. The explanatory strength of Norton's identity investment framework and Dörnyei's Ideal L2 Self construct is thus validated within the Vietnamese Gen Z context. Participants did not merely want to learn English; they wanted to become the kind of person who speaks English fluently, confidently, and naturally, a vision of the self that is simultaneously aspirational and motivationally generative.

The vocational context of COFER adds a dimension that Thu and Lee's (2024) L2MSS research did not fully address: Ideal L2 Self orientations appear to operate motivationally not only among language majors but across disciplines such as logistics, e-commerce, and business administration. This suggests a broader relevance of identity-based motivation than previously documented in Vietnamese undergraduate settings. Darvin and Norton's (2015) claim that digital platforms have become primary sites for identity investment was borne out clearly in the data. Several participants described watching TikTok creators their own age speak English effortlessly and feeling simultaneously inspired and unsettled. That tension fed directly into their motivation to close the gap between who they are now and who they wish to become.

5.3. Digital Media as a Transformative Motivational Environment

A key finding is that a large majority of participants described English as feeling more accessible and motivating through social media and AI, identifying digital media as a significant source of intrinsic motivation. This finding supports and expands Sockett's (2014) Online Informal Learning of English framework. TikTok, YouTube, music streaming, online gaming, and AI conversational tools are demonstrated to function as motivational ecosystems where intrinsic engagement with English is sustained, rather than mere supplementary resources.

Particularly noteworthy is the role of AI tools, especially ChatGPT, as low-anxiety communicative practice environments. Participants who described using AI as a conversational partner were effectively creating conditions that MacIntyre et al. (1998) identify as optimal for WTC development: low evaluative threat, high communicative control, and personally meaningful interaction. This suggests that AI-mediated English practice may serve as a critical bridge between participants' high intrinsic motivation and their low willingness to communicate in face-to-face contexts a pedagogical implication with significant practical relevance.

5.4. The Motivational-Communicative Gap: A Critical Pedagogical Challenge

The principal innovative contribution of this research lies in the identification and characterization of a motivational-communicative disparity. A systematic disconnect between participants' inherent drive to interact with English and their limited readiness to communicate in genuine, interpersonal scenarios is highlighted. This disparity, vividly illustrated in participants' figurative self-portrayals, implies that inherent motivation, while essential, is insufficient to generate communicative behavior when foreign language anxiety serves as a conflicting emotional influence.

This finding corroborated and expands upon existing willingness-to-communicate research. Peng's (2014) demonstration that classroom anxiety supersedes motivational elements in interpersonal communicative situations is mirrored here. However, an important refinement is added by the current study: the motivational-communicative disparity is not perceived as an unyielding obstacle but as a dynamic, emotionally intricate arena. Within this space, shame, anxiety, and inherent motivation interact in manners that can, under propitious circumstances, yield renewed motivational involvement.

Several participants described how humiliating communicative setbacks had paradoxically revitalized their self-guided learning. This suggests that the disparity, while problematic, is not motivationally inactive. This dynamic is consistent with Cao & Philp’s (2006) finding that willingness to communicate fluctuates significantly across contexts. It is suggested that pedagogical interventions aimed at diminishing evaluative threat through collaborative, low-stakes speaking exercises, AI-mediated practice, and identity-affirming classroom environments may prove more efficacious in bridging the motivational-communicative disparity than conventional communicative language teaching methodologies that prioritize accuracy and formal evaluation.

5.5. Vietnam's ESL Policy Shift and Its Motivational Implications

Vietnamese students have shown significant interest in updated ESL instruction policies. When asked directly what they made of the redesignation, participants most often described it as raising the stakes attached to English: it would cease to be a competitive advantage and become a condition of ordinary participation in study and work. Many said this would raise their motivation, although several also named pressure, and a recurring concern

was the fear of being left behind. English is no longer viewed as just a subject studied for exams. Instead, it is seen as a powerful tool for achieving personal development and expertise.

Participants discussed the redesignation in terms of opportunity, requirement and pressure rather than in terms of classroom practice. Ushioda (2011) cautions against superficial adherence to top-down policy directives. Fundamental pedagogical change may not be elicited solely through institutional acknowledgment. For optimal motivational impact, policy revisions must be implemented alongside modified assessment methods. Teacher training and classroom practices should also be adjusted accordingly.

For learners, a supportive educational setting is ultimately required, instead of merely imposed regulations. In this environment, autonomy should be cultivated, and communicative experimentation should be embraced. Individual expression must be validated. Otherwise, improvements, even those well-intended, risk being superficial. Their presence is only noted in formal documentation, and their fulfillment in real classroom settings is unmet.

Whether the 2023 redesignation ultimately reshapes classroom practice remains to be seen. What the data make clear, however, is that the motivational groundwork among Gen Z learners at COFER already exists. These students are not waiting to be motivated by policy; they arrived with it. The more pressing question for educators and institutions is not how to spark engagement, but whether the learning environment is structured in ways that allow that engagement to surface, grow, and translate into communicative confidence.

6. Conclusion

Vietnam’s 2023 shift toward an ESL framework prompted a closer look at what truly motivates the current generation of learners. By analyzing 83 COFER students, this research highlights five distinct areas: enjoyment-based engagement, identity investment, reliance on digital media, autonomous learning, and a visible gap between feeling motivated and speaking out. These results challenge the traditional view of Vietnamese students as purely instrumentally driven. Rather than focusing on grades or institutional policies, the participants showed a strong desire to learn based on their future self-image.

The study makes three principal contributions. Initially, the prevailing depiction of Vietnamese EFL learners as instrumentally driven is contested. The presence, and indeed dominance, of intrinsic orientations among Gen Z students at COFER is showcased, as understood through SDT's motivational spectrum. Second, it establishes digital media spaces as autonomous motivational ecosystems, not merely supplementary tools, extending OILE and identity investment frameworks into the Vietnamese Gen Z context. Subsequently, the motivational-communicative gap is pinpointed and theorized as a unique attribute of this learner group's motivational profile. Notable implications for research and pedagogy are suggested.

Several limitations of this study warrant acknowledgment. First, the sample was drawn exclusively from one institution, COFER, which serves a predominantly urban, career-oriented student population in Ho Chi Minh City. As such, the findings may not be fully transferable to learners in rural areas, different institutional contexts, or other Vietnamese universities with distinct pedagogical cultures. There are also methodological constraints to consider. The reliance on self-reported written interviews provided deep personal insights, but participants might still have tailored their answers, especially knowing the researcher is a faculty member at the same institution. Furthermore, this research represents a cross-sectional snapshot taken in March 2026. It cannot measure how student motivation might fluctuate across different academic years or adapt as the new ESL policy takes deeper effect. To build on this work,

subsequent studies should consider longitudinal tracking, mixed methods, or expanding the sample to other universities. This would help confirm how widely these motivational trends apply across Vietnam's changing educational landscape.

From a pedagogical perspective, the incorporation of students' digital activities into curriculum design is necessitated by these findings. Furthermore, AI should be leveraged as a supportive space for communication. Assessment cultures that value experimentation and acknowledge learner identity, rather than strict evaluation, ought to be prioritized. Forthcoming research endeavors should implement longitudinal, multi-site, or mixed-methods approaches. Motivational development throughout university can then be observed. Furthermore, interventions for the motivational-communicative gap should be evaluated.

The intrinsic motivational resources that Gen Z brings to English learning already exist within them. The challenge for educators and policymakers is not to manufacture motivation from nothing, but to recognize, appreciate, and actively cultivate the intrinsic resources that these learners carry into every classroom, every digital screen, and every encounter with the English language.

Reference

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa.

Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System, 34*(4), 480-493. https://doi.org/10.1016/j.system.2006.05.002.

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches. SAGE Publications.

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics, 35*, 36-56. https://doi.org/10.1017/S0267190514000191.

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Deci, E., & Ryan, R. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01.

Dörnyei, Z. (2009). 2. the L2 motivational self system. In Motivation, Language Identity and the L2 Self (pp. 9-42). Multilingual Matters. https://doi.org/10.21832/9781847691293-003

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation.* Pearson Education.

Dressman, M., & Sadler, R. W. (Eds.). (2019). The handbook of informal language learning. Wiley.

Gardner, R. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* Edward Arnold.

Geertz, C. (1973). *The interpretation of cultures: Selected essays.* Basic Books.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x.

Lai, C. (2017). Autonomous language learning with technology: beyond the classroom.

Le, M. T. T., & Nguyen, T. H. (2023). Motivation of EFL Vietnamese students in economics-related majors. *Vietnam Journal of Education, 7*(3), 263–275. https://doi.org/10.52296/vje.2023.329.

MacIntyre, P., Clément, R., Dörnyei, Z., & Noels, K. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x.

Nguyen, T. P. L. (2022). Language learning motivation of Vietnamese EFL students and their attitudes towards classroom factors. *Ho Chi Minh City Open University Journal of Science – Social Sciences, 12*(1), 63–78. https://doi.org/10.46223/HCMCOUJS.soci.en.12.1.2282.2022.

Nguyen, V. S., & Habók, A. (2021). Vietnamese non-English-major students' motivation to learn English: From activity theory perspective. *Heliyon, 7*(4), e06819. https://doi.org/10.1016/j.heliyon.2021.e06819.

Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change.* Longman.

Norton, B. (2013). *Identity and language learning: Extending the conversation* (2nd ed.). Multilingual Matters. https://doi.org/10.21832/9781783090563.

Peng, J.-E. (2014). *Willingness to communicate in the Chinese EFL university classroom: An ecological perspective.* Multilingual Matters. https://doi.org/10.21832/9781783091560.

Rahaman, A. (2025). Identity, ideology, and capital in EFL learning: A narrative case study. *Asian-Pacific Journal of Second and Foreign Language Education, 10*(1), 13. https://doi.org/10.1186/s40862-024-00315-8.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68.

Seemiller, C., & Grace, M. (2016). *Generation Z goes to college.* Jossey-Bass.

Sockett, G. (2014). *The online informal learning of English.* Palgrave Macmillan.

Thu, N. T. T., & Lee, M. W. (2024). L2 motivational self system and its impact on English learning efforts in Vietnamese EFL context. *The Asia-Pacific Education Researcher, 34*(4), 1375–1384. https://doi.org/10.1007/s40299-024-00951-1.

Tran, L. T. (2007). Learners' motivation and identity in the Vietnamese EFL writing classroom. *English Teaching: Practice and Critique, 6*(1), 151–163.

Ushioda, E. (2011). Why autonomy? Insights from motivation theory and research. *Innovation in Language Learning and Teaching, 5*(2), 221–232. https://doi.org/10.1080/17501229.2011.577536.

Bionote

M.A. Le Nguyen Hoang Nam is currently lecturing at the Faculty of Foreign Languages, College of Foreign Economic Relations (COFER), in Ho Chi Minh City, Vietnam. Second language motivation is at the core of his research, coupled with Gen Z learner identity. Digital informal learning and the evolving ESL context within Vietnam are also explored. Qualitative and identity-based frameworks are employed in his work. The shaping of intrinsic motivation and autonomous learning within Vietnamese university students through digital environments is examined.

#110 AI-SUPPORTED WRITING SCAFFOLDING AND PERCEPTIONS OF ACADEMIC INTEGRITY AMONG EFL JUNIORS AT HUFLIT

Anh Tuan Nguyen Le[1], Ngoc Huyen Thai Thi[2], Minh Chau Diep[3]

Abstract

This research explores how EFL Juniors at HUFLIT view academic integrity as they utilise AI-assisted writing scaffolds for their academic writing course (Introduction to Research Methods – Project 2). Generative artificial intelligence tools have been integrated into institutions of higher education; however, this has also created a conflict over whether there is a benefit or detriment to using these tools for educational purposes and to promote the authenticity of authorship. Data were collected through questionnaires, reflective papers, and semi-structured interviews. Descriptive statistics and thematic analysis were used to identify patterns of AI use, perceptions of academic integrity, and recommended practices for responsible AI use. The results indicate that most EFL Juniors see AI as an advantageously useful writing scaffold tool in terms of helping them create ideas, organise content and enhance vocabulary, ultimately creating greater confidence in their ability to write well. Nonetheless, caution needs to be taken regarding the possibility of too much dependence upon AI, and the effect it could have on students' writing skills. In addition, this research provides a practical understanding of how to integrate AI functionality into their teaching practices to develop their EFL Juniors' writing skills while remaining true to the principles of academic integrity.

***Keywords*:** AI-assisted writing, Project 2, EFL Juniors, mixed-methods approach, teaching practices

1. Introduction

1.1. Study background

The higher education landscape as we know it today is likely to undergo a radical transformation by 2025. This transformation will be driven by the deployment of advanced AI tools in the capacity of always-on online tutors and allocation analytics tools that provide personalised feedback to the learners (Kestin et al., 2025). The use of AI tools will allow for processing and generating content in a variety of formats, from text, code, and images, which in turn, will provide multiple ways of meeting the linguistic needs of learners from different backgrounds, such as those of multilingual communities who need help creating a sophisticated language base (Nyaaba, 2023). According to Quad C (2025), more than 80% of all university students around the globe have used some type of GenAI to help themselves learn through summarising materials, generating concepts, and even outlines; consequently, AI literacy will be

[1]*Corresponding author, Bachelor of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology, HCMC, Vietnam; Email: tuannguyen03012004@gmail.com*
[2]*Bachelor of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology, HCMC, Vietnam; Email: thaithingochuyen.470dt@gmail.com*
[3]*Lecturer, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology, HCMC, Vietnam; Email: chaudm@huflit.edu.vn*

an essential skill to get hired nowadays. As powerful as AI technologies are when it comes to assisting people in writing and researching, there were multiple discussions about how the rise in usage of such software jeopardises academic integrity and makes it hard to discern what is considered original work by humans and not generated by algorithms. Despite having beneficial aspects in writing, such as taking care of most of the mental processes involved and giving confidence to authors of any works, AI technologies bring risks associated with over-relying on them and producing misleading or non-existent references to the literature (Salleh et al., 2025). Thus, the aim here would be to have an open discussion between us as educators and our students concerning the values of academic integrity.

1.2. Problem statement

The employment of AI generational tools can also help students structure their research paper for the very first time. Indeed, according to Project 2 in the Introduction to Research Methods course at HUFLIT, AI tools show that one of the main problems in the process of writing is connected with maintaining the integrity of their papers while benefiting from the support provided by the AI technology. The EFL Juniors have faced the problem of cognitive overload since they should understand research terminology and use precise linguistic structures simultaneously. As a result, many of the EFL Juniors turn to AI tools for various purposes, including generating ideas, structuring the outline, and polishing language. Nevertheless, it is important to admit that there is a great difference between the technical skills of working with these tools and the overall knowledge of the capabilities of AI. In other words, the students can hardly distinguish the inherent ethical limitations of writing and be critical about the content produced with the help of AI. There are also several preliminary findings related to the fact that, despite the fact that the students believe that AI will motivate them and help them to cope with writing anxiety, there is no established framework regulating the process of applying this tool in the university. Moreover, there is a high risk of students relying too much on AI and, therefore, handing their teachers their papers without giving any credit for the information taken from them. Thus, without knowing whether the students perceive AI as scaffolding or a violation of academic integrity, the professor cannot be sure about the proper educational process.

1.3. Research objectives

The main purpose of this study is to analyse systematically how the subjective experience and behaviours of EFL Juniors using AI tools affect their role as cognitive and linguistic helpers in doing academic writing tasks at HUFLIT. The current study focuses on the attitudes of students toward the capability of AI in fostering rather than compromising the integrity principles, as well as assisting in maintaining balance between appropriate and inappropriate use of AI to avoid "*AI-giarism*". Furthermore, the current study explores the particular steps they take while using AI, including brainstorming topics for research papers, changing sentences to make them better in English, making sure the reasoning behind them is sound, or even coming up with a first draft, which they revise afterwards.

1.4. Delimitations

The research is only conducted among EFL Juniors at HUFLIT who are enrolled in the Introduction to Research Methods (Project 2) class. About the technological scope of the inquiry, the study considers only the use of AI-based writing assistance such as brainstorming, structuring,

and text editing, without considering AI tools for speaking, listening, and reading skills. Moreover, the scope of the research will be restricted to an examination of the perception of academic integrity and AI assistance among undergraduates. This means that the empirical efficiency of AI tools about writing, language proficiency, and academic performance will not be evaluated. Finally, the scope of analysis of the research will be confined to a descriptive mixed-methods approach, whereby only descriptive statistics will be used for quantitative data, while inductive thematic analysis will be used for qualitative interview transcripts.

1.5. Research questions

In response to these research purposes, the current study is expected to answer the following three research questions:

1. In what way do EFL Juniors employ AI tools in scaffolding Project 2 academic writing tasks?

2. What are the perceptions about the academic integrity of EFL Juniors when using AI tools for Project 2 academic writing?

3. What strategies could be proposed to foster academic integrity?

2. Literature Review

2.1. Theoretical framework

Vygotsky (1978) conceptualises the Zone of Proximal Development (ZPD) as the difference between a learner's independent capacity and their potential achievement under expert guidance. The concept provides the theoretical basis for instructional scaffolding. In the case of EFL Juniors, AI tools function as digital scaffolds by providing real-time feedback and adaptive prompts. These interventions assist students in managing demanding academic tasks, such as the organisation of research methods, thereby facilitating performance that exceeds their current independent level (Nordlof, 2014). Recent studies suggest that AI integration within the ZPD framework can improve writing skills through personalised assistance and the promotion of self-regulation, while remaining consistent with standards of academic integrity (Gayed et al., 2025).

The Technology Acceptance Model (TAM), proposed by Davis (1989), posits that technology adoption is primarily influenced by perceived usefulness and perceived ease of use. Perceived usefulness refers to the user's belief that a system will enhance their performance, while perceived ease of use concerns the effort required to use the system. Within AI-assisted writing, perceived usefulness encompasses student beliefs regarding improvements in idea generation, text coherence, and grammatical accuracy. Perceived ease of use reflects the extent to which students find the interface and operational requirements of these tools manageable. While the TAM has been refined over time, recent applications examine the relationship between technology acceptance and student perceptions of academic integrity.

2.2. AI-Supported Scaffolding in Academic Writing

The integration of AI tools has significantly changed how academic writing is approached, primarily by offering structured scaffolding that assists students at various stages of the process. Generative models like ChatGPT are notable for providing immediate feedback, which helps

students brainstorm and organise their thoughts more effectively (Khalifa & Albadawy, 2024). These applications now serve critical functions across several research dimensions, including idea generation, content structuring, literature synthesis, and data management, alongside improvements in communication and ethical editing (Khalifa & Albadaway, 2024). Despite their educational benefits, the growing prevalence of AI also presents substantial risks to academic integrity, particularly when students lack clear ethical frameworks or institutional guidance. One prominent concern is the possibility of undetectable plagiarism, as AI systems produce fluid, original-appearing text that often bypasses traditional detection tools (Fowler, 2023). Furthermore, an over-reliance on these technologies may compromise critical thinking and independent writing skills, as students may prioritise automated content over the deep cognitive engagement necessary for academic growth (Richards et al., 2024). Finally, the lack of transparency in AI-generated work poses a challenge for assessment, making it harder for educators to evaluate the actual effort and knowledge a student has acquired (Bjelobaba et al., 2024).

2.3. Academic Integrity in AI-Supported Writing

The rise of AI-assisted plagiarism has become a pressing issue within the academic community (Chan, 2023). This practice, which involves submitting AI-generated or heavily modified content without proper disclosure, has significantly blurred the distinction between independent authorship and automated assistance. As generative models grow more sophisticated, their ability to produce coherent, contextually appropriate academic prose makes them increasingly difficult to distinguish from human work, thereby challenging long-standing definitions of plagiarism (Torres-Diaz et al., 2025). For EFL Juniors, the concept of academic integrity in an AI-supported environment is often nuanced and occasionally contradictory, reflecting a balance between pedagogical utility and ethical responsibility. Many of these students have begun to view integrity not merely as the absence of plagiarism, but as a broader commitment to transparent practice, original thought, and intentional learning. Research suggests that these learners generally regard AI as a valuable asset for linguistic development, particularly for refining grammar, broadening vocabulary, and facilitating the initial stages of idea generation.

2.4. Pedagogical Strategies for Maintaining Academic Integrity

Research suggests that while AI tools can improve grammatical accuracy and help structure early drafts, integrating them without clear pedagogical guidance risks fostering an over-reliance on automated outputs. This dependency may, in turn, impede the development of higher-order writing skills, particularly critical analysis and complex argumentation (Aljuaid, 2024). To counter this, educators are encouraged to adopt process-oriented frameworks that require students to document their use of AI, making their creative choices more transparent and accountable. At the same time, institutional clarity is vital; vague or inconsistent standards often result in unintentional academic misconduct (Bui & Tong, 2025). Guidelines must therefore explicitly distinguish between acceptable assistance and prohibited practices, such as presenting AI-generated text as one's own. This is especially important given that AI-driven paraphrasing can easily obscure the origins of an idea, raising the risk of plagiarism when sources are not properly cited (Hernandez et al., 2024).

3. Methodology

3.1. The participants

The participants in this study were 176 EFL Juniors enrolled in the Introduction to Research Methods (Project 2) course at HUFLIT during the 2025–2026 academic year. The cohort represents a variety of academic specialisations. The selection of this specific course was based on its role in the undergraduate curriculum, which marks the transition from initial data collection in Project 1 to the development of more complex academic discourse in Project 2. In particular, data were collected through a questionnaire with various question types and in-depth interviews to obtain data that could provide rich insights into these findings.

3.2. Research Design

This study employs a descriptive mixed-methods design and is guided by research questions. The aim is to describe students' use of AI, perceptions of academic integrity, and practical recommendations, rather than to examine causal relationships between variables. A mixed-methods approach was employed, combining descriptive statistics and thematic analysis to provide a triangulated empirical view of this research. Specifically, these materials were meticulously designed to ensure methodological solutions, ethical compliance, and, ultimately, the research goal that aims to examine AI-supported scaffolds and learning motivation towards undesirable behaviours such as dishonesty relevant to academic integrity among EFL Juniors at HUFLIT.

3.3. Instruments

3.3.1. Questionnaire

The primary instrument was a 17-item online questionnaire for EFL Juniors. The questionnaire was intended to collect EFL Juniors' perceptions of AI-based learning assistance associated with academic integrity in Project 2. All 17 questions were split into four sections, corresponding to the objectives of the study and the three research questions.

- Section one, Background Information (Questions 1 to 3), was focused on collecting the respondents' gender information, frequency of AI tool application, and familiarity with AI learning assistance.
- Section two, AI usage to scaffold Project 2 writing (Questions 4 to 6), explored particular types of AI assistance, its impact on particular aspects of writing, and associated advantages and potential risks.
- Section three, perceptions of academic integrity, including citation awareness and plagiarism awareness when applying AI tools to Project 2 writing (Questions 7 to 14) via the commonly-used 1-to-5 scale.
- Finally, the fourth section (Questions 15 to 17) contained one multiple-choice question on EFL Juniors' suggestions for maintaining academic integrity while employing AI tools, in addition to two open-ended ones encouraging students to share more insights about responsible utilisation of AI tools.

3.3.2. In-depth interview

In-depth interviews were conducted with 6 EFL Juniors (each participant lasted about 10 minutes) to enrich the findings obtained through surveys with the help of EFL Juniors with practical and detailed perspectives on their perceptions of academic integrity in an AI-based context:

- Section 1 (Research Question 1): Questions 1 and 2 address the manner in which EFL Juniors employ AI tools to scaffold particular Project 2 academic writing assignments, along with any challenges involved in this process.
- Section 2 (Research Question 2): Questions 3, 4, and 5 aim to examine students' perceptions of academic integrity, namely, how they manage the balance between personal writing and AI assistance and establish ethical boundaries.
- Section 3 (Research Question 3): Questions 6 and 7 provide information on students' recommendations regarding practical guidelines for upholding academic integrity.

3.4. Data analysis

For an in-depth analysis of the data collected, both descriptive statistics and thematic analysis were used and connected with the main research questions. For the process of analysis of the data in the descriptive statistics, IBM SPSS Statistics was used to analyse the survey data. Variables were recoded using numbers (0 – No; 1 – Yes) for demographic variables and multiple response items to be able to measure frequency and percentages to define popular types of application use and their distribution under RQ1. Additionally, the data obtained on the Likert scale, both from the structured matrix and open-ended questionnaires, were analysed via the mean and standard deviation measures in order to assess the level of students' ethical awareness (RQ2) and preferences (RQ3). In connection with that, the data from the open-ended questionnaire and the semi-structured interviews were subjected to the inductive thematic analysis. The guide for conducting the semi-structured interviews was based on the concepts of academic integrity and technological literacy developed by McCabe (2005) and Selber (2004). It was fully translated into English and then transcribed. Qualitative data were classified into several themes in order to determine deeper perceptions of students about over-reliance (RQ2) and their curriculum suggestions (RQ3), keeping their anonymity intact (P1-P6).

4. Results

4.1. Demographic information

Table 1

Survey Respondents' Demographic and Behavioural Statistics

		1. Gender	2. How often do you use these AI tools for Project 2 writing?	3. What is your familiarity level with the concept of "AI-supported scaffolding" in academic writing?
N	Valid	176	176	176
	Missing	0	0	0

Table 2

Frequency Distribution of Respondents by Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
alid	Female	121	68.8	68.8	68.8
	Male	55	31.3	31.3	100.0
	Total	176	100.0	100.0	

The demographic characteristics of the 176 students who took part in this study have been outlined in Table 1. No data were missing in the analysis because all respondents provided answers to all questions asked. Speaking about gender balance within this group of students in Table 2, one can note that females made up to 68.8%, whereas males made up 31.3%. Such a distribution of genders is characteristic of the demographic data of students studying in language departments of universities.

4.2. EFL Juniors' employment of AI tools

Table 3

Frequency of AI Tools Usage in Project 2

		Frequency	Percent	Valid Percent	Cumulative Percent
alid	Always (weekly or more)	60	34.1	34.1	34.1
	Never	1	0.6	0.6	34.7
	Often (once a month or more)	52	29.5	29.5	64.2
	Rarely (once a semester or less)	8	4.5	4.5	68.8
	Sometimes (a few times per semester)	55	31.3	31.3	100.0
	Total	176	100.0	100.0	

The current data points out the high popularity of the use of AI tools among Project 2 students. According to Table 3, the percentage of those who stated to "*always use*" such technologies was equal to 34.1%. Furthermore, 29.5% of respondents said that they used AI tools "*often*", meaning that they used them once a month or even more often. In turn, another 31.3% said that they "*sometimes*" used the discussed technology – a few times per semester. Only 4.5% of students stated to use these technologies "*rarely*" and just 0.6% stated "*never*". Thus, it becomes clear that more than 94% of students use AI.

Anonymously labelled as P1-P6, student testimonies indicate a careful matching of particular technologies to individual assignments. At the initial cognitive phase, students use

conversational engines to collaborate actively in their processes. As corroborated by P1, "*my team worked with several AI tools like ChatGPT, Gemini, Deepseek, and also NotebookLM. We also use Grammarly… ChatGPT and Gemini are kind of common, and we mainly use them for brainstorming"*. In this respect, P6 supports P1's assertion, claiming that "*actually, I use AI as a brainstorming partner. Instead of a blank space, I use it to break down the complex requirements of the project*". Once the process of writing unfolds, students turn to using technologies for formatting ideas, changing vocabulary, and polishing the language. For instance, according to P3, "*the main role of AI is to help me in arranging and arranging of ideas... After I used AI to help me, I think I know how to write better, as well as how to write more clearly*". Similarly, P4 regularly applied precise prompts to external databases, revealing, "*I will give AI a topic... and then I will ask AI to help me connect complex structures to make my research more logical... It will give me more complex words and sentences... to help me write more logically*". Lastly, for superficial corrections, students used special editing engines to avoid technical errors. Thus, P2 emphasized, "*I would use my AI to organise my ideas and for revising graphs, because I'm always making spelling mistakes*". Likewise, P5 highlighted, "*I often use AI specifically to see if I make any mistakes... or to correct the grammatical errors*". In sum, these testimonies illustrate that undergraduates widely apply various AI technologies in order to organize, paraphrase, and improve their assignment

Table 4

Types and Frequency Distribution of AI Tools Utilized by EFL Juniors

		Responses		Percent of Cases
		N	Percent	
AI tools used by EFL Juniors[a]	ChatGPT	159	30.9%	90.3%
	Gemini	143	27.8%	81.3%
	Quillbot	90	17.5%	51.1%
	Copilot	40	7.8%	22.7%
	Google Translate	70	13.6%	39.8%
	ZeroGPT	1	0.2%	0.6%
	Grammarly	2	0.4%	1.1%
	Grok	4	0.8%	2.3%
	Claude	3	0.6%	1.7%
	NotebookLM	1	0.2%	0.6%
	Perplexity	2	0.4%	1.1%

	Total	515	100.0%	292.6%

From Table 4, it is evident that the students rely heavily on generative AI tools. The most preferred tools among students include ChatGPT, used by 90.3% of them, and Gemini, used by 81.3%. Other commonly used paraphrasing and translation tools include Quillbot (used by 51.1% of students) and Google Translate (used by 39.8%). Other less common but still popular AI tools among the students include Copilot, with 22.7% usage. On the other hand, some specialised AI tools like Grok (only used by 2.3% of students), Claude (used by 1.7% of them), Grammarly (by 1.1%), Perplexity (used by 1.1%), NotebookLM (0.6%) and ZeroGPT (0.6%) remain infrequently used by the group.

Anonymised as P1 to P6 for privacy reasons, the transcripts of the interviews show that conversational models and text rewriter platforms are the most visited among the students' research routines. In terms of high-frequency brainstorming and conceptual setup, conversational platforms like ChatGPT and Gemini act as the main foundations of the research activities. According to P1, the following explanation was offered: "*my team worked with several AI tools like ChatGPT, Gemini, Deepseek, and also NotebookLM. We also use Grammarly, and each one supports a different part of our writing process. ChatGPT and Gemini are kind of common, and we mainly use them for brainstorming"*. This heavy dependence on particular software brands for structuring layouts and draft creation is a frequent trend among multiple individual research processes. For instance, P2 admitted, "*I usually use Chattivity [ChatGPT] to do it... Sometimes I feel burnt out sometimes. So sometimes I just throw it on AI to write it for me*". Also, P4 described their systematic distribution of instructions across engines and conversational rewriters during complex task execution: "*I will give AI a topic that I want... and then I will ask AI to help me connect complex structures to make my research more logical*". Lastly, for simple mechanical proofreading and polishment of a draft, automation editors become frequently applied in the final stage. As stated by P1, "*Grammarly to check grammar, punctuation, and sentence clarity*". At the same time, both P5 and P6 relied extensively on the software for similar functions. In particular, P6 added, "*I use it for, um, finding those advanced vocab for my paragraph*".

Table 5

EFL Juniors' Evaluation of AI Features for Project 2 writing

		Responses		Percent of Cases
		N	Percent	
AI features evaluated by EFL Juniors[a]	Grammar and style correction	134	38.0%	77.5%
	Paraphrasing existing text	89	25.2%	51.4%
	Content idea brainstorming	90	25.5%	52.0%
	Reference and citation formatting	40	11.3%	23.1%
	Total	353	100.0%	204.0%

a. Dichotomy group tabulated at value 1.

As per the detailed feature usage statistics in Table 5, "*Grammar and style correction*" is the most frequently selected function, being chosen by 77.5% of the participants (i.e., comprising 38.0% of all feature usage responses). Besides surface-level edits, undergraduates extensively use cognitive features to support their text creation process. Namely, 52.0% of the students use "*Content idea brainstorming*", while another highly selected cognitive feature is "*Paraphrasing existing text*" at 51.4%. At the same time, fewer students utilise technical features in the context of academic writing management, since "*Reference and citation formatting*" is chosen by 23.1% of the respondents.

In the in-depth interview, P1 and P6 focused their review of the examined platforms primarily on high-level cognitive and brainstorming features, using conversation-based tools as interactive assistants to break down the complexity of required research structures. Thus, P1 found the examined services to be highly efficient in assisting users with "*brainstorming, outlining, grammar checking, and revision*", while P6 explained their use of the mentioned platforms, saying, "*Actually, I use AI as a brainstorming partner. Instead of a blank space, I use it to break down the complex requirements of the project, too. Um, and this would actually help me a lot, uh, for the paragraph cohesion*". Further discussing structural layout and synthesis functions, P3 and P4 viewed the organisational capabilities of AI features positively, as they helped them with organising "*the arrangement and arranging of ideas... after I used AI to help me, I think I know how to write better, as well as how to write more clearly*". On the other hand, P4 valued the connection-making capabilities of AI features as follows: "*I will ask AI to help me connect complex structures to make my research more logical... It will give me more concise sentences... as well as more complex words and sentences*". Finally, concerning language-polishing and text-refining features, P2 and P5 were looking for capabilities that eliminated surface-level errors from texts. As P2 stated, "*The one thing that makes me choose AI software is revising graphs, because I'm always making spelling mistakes*", while P5 stated that they used AI-based tools "*specifically to see if I make any mistakes... [and] to correct the grammatical errors, or to correct the weird sentences*".

Table 6

Evaluation of AI Benefits of AI-Supported Writing

		Responses		Percent of Cases
		N	Percent	
AI benefits evaluated by EFL Juniors[a]	Improved writing structure	106	20.7%	60.2%
	Quick grammar error reduction	116	22.7%	65.9%
	Creative idea suggestions	127	24.9%	72.2%
	Time savings in research	113	22.1%	64.2%
	Increased writing confidence	49	9.6%	27.8%
	Total	511	100.0%	290.3%

Based on the specific benefits provided in the table above, cognitive and creativity help is the first benefit mentioned by most respondents. Particularly, the highest percentage of people

mentioned "*Creative idea suggestions*" as the main advantage, with 72.2% (24.9% overall response percentage) mentioning this benefit specifically. Practical and structural help are also among the best advantages identified by the participants. Thus, the second highest rate was "*Quick grammar error reduction*" (65.9%), followed by "*Time savings in research*" (64.2%), and "*Improved writing structure*" (60.2%). At the same time, psychological help is relatively less widespread; only 27.8% chose "*Increased writing confidence*".

For the in-depth interview, in terms of cognitive development, it is clear that participants noted that communication with AI helped them diversify their views on the subject and improve their analysis of particular issues. Thus, P1 noted that comparison of the two models helped him obtain mental clarity since platforms like DeepSeek became useful for "*if I wanted another perspective or a different explanation of some point*". P2 obtained similar cognitive help when he was learning to navigate a quantitative tool; thus, according to him, AI helped because "*I'm still using AI to learn how to use the analyser tool like SPSS... I have learned a way to use AI more effectively*". For language-driven enhancement, software served as a good linguistic scaffold that improved the academic quality of the text. For example, P1 stated that he was able to use "*more advanced vocabulary which makes me sound more academic*", whereas P4 enjoyed "*more concise sentences... as well as more complex words... to help me write more logical sentences*". Moreover, P6 mentioned that the analysis of the AI's work helped him express himself better: "*By seeing how AI can help me to rephrase... I learned how to express my ideas*". Finally, in terms of workflow optimisation, AI was seen as convenient and efficient enough for time savings and structural improvements. P3 noticed that AI helped him improve both his skills and writing, making him "*know how to write better, as well as how to write more clearly*"; P4 agreed with this statement but focused on the time aspect: "*It will give me quick answers, as well as save me time*".

4.3. EFL Juniors' Perceptions of Academic Integrity

Table 7

Perceptions of Academic Integrity Awareness when Using AI Tools

	Mean	Std. Deviation	N
AI tools increase my awareness of proper citations	3.84	.945	176
AI tools helped me distinguish between acceptable AI assistance and academic cheating.	3.71	.895	176
AI tools raise my awareness of plagiarism risks	3.97	.871	176

Table 8

Reliability Statistics for the Scale

Cronbach's Alpha	N of Items
.637	3

Table 9

Cronbach's Alpha Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
AI tools increase my awareness of proper citations	7.68	2.150	.440	.548
AI tools helped me distinguish between acceptable AI assistance and academic cheating.	7.81	2.305	.425	.567
AI tools raise my awareness of plagiarism risks	7.55	2.261	.474	.501

Descriptive measures in Table 7 show that there is high student agreement regarding the effectiveness of AI as a means to enhance ethical awareness. Indeed, the last item demonstrates the highest level of consensus among the group members: Students strongly agree that AI increases their awareness of plagiarism risks (M = 3.97, SD = .871). This result is followed by the item that AI boosts students' awareness of citing practices (M = 3.84, SD = .945) and helps to differentiate between permissible and forbidden types of assistance (M = 3.71, SD = .895). In addition, the overall internal consistency of this 3-item sub-scale is verified in Table 8, returning a Cronbach's Alpha value of .637. While moderate, this meets acceptable research thresholds for a concise behavioural sub-scale. Finally, the item-total statistics in Table 9 prove that all three variables correlate well with the overall scale, as corrected item-total correlations remain safely above .40. Removing any of the items would lower the overall Cronbach's Alpha to values between .501 and .567, confirming that keeping all three statements provides the most reliable measurement framework.

Table 10

Academic Integrity's Risks and Challenges

	Mean	Std. Deviation	N
AI for editing or paraphrasing in Project 2 is a violation of academic integrity	3.39	1.036	176
Over-reliance on AI tools may reduce my independent writing skills	3.99	.881	176
It is sometimes difficult to distinguish between my own ideas and AI-generated content.	3.82	.943	176
Reliance on AI tools for scaffolding in Project 2 increases the risk of unintentional academic dishonesty.	3.96	.824	176
Using AI-supported scaffolding in Project 2 provides an unfair advantage over peers who do not use it, affecting academic integrity.	3.73	1.004	176

Table 11

Reliability Statistics for the Scale

Cronbach's Alpha	N of Items
.686	5

Table 12

Cronbach's Alpha Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
AI for editing or paraphrasing in Project 2 is a violation of academic integrity	15.51	6.571	.409	.653
Over-reliance on AI tools may reduce my independent writing skills	14.91	7.615	.293	.694
It is sometimes difficult to distinguish between my own ideas and AI-generated content.	15.07	6.823	.428	.642
Reliance on AI tools for scaffolding in Project 2 increases the risk of unintentional academic dishonesty.	14.94	6.882	.522	.607
Using AI-supported scaffolding in Project 2 provides an unfair advantage over peers who do not use it, affecting academic integrity.	15.16	5.990	.574	.572

In Table 10, as for the hidden risks associated with the use of AI writing aids, the participants demonstrate that they are very much aware of this factor. In this regard, participants show the highest level of agreement with the item related to concern that excessive reliance on AI tools may hinder the development of students' independent writing skills (M = 3.99, SD = .881). In addition, the statement that reliance on scaffolding provided by AI tools can contribute to unintentional academic dishonesty (M = 3.96, SD = .824). Furthermore, it appears to be fairly difficult for students to distinguish between original ideas and AI-generated content (M = 3.82, SD = .943). Moreover, it is worth mentioning that participants agree that using AI tools provides a competitive advantage in relation to peers who do not employ such means (M = 3.73, SD = 1.004). At the same time, the item with the lowest level of student agreement indicates that there are doubts whether editing and paraphrasing constitute academic dishonesty (M = 3.39, SD = 1.036).

The overall internal consistency for this 5-item sub-scale is verified in Table 11, which shows a Cronbach's Alpha value of .686. This coefficient indicates an acceptable level of reliability for an undergraduate perceptual survey, proving that the grouped questions reliably measure the negative ethical concerns and academic anxieties felt by students. Looking deeper into the scale dynamics, the item-total metrics in Table 12 show a generally solid relationship among the items, except for the second item, which displays a lower corrected item-total

correlation of .293. Despite this weak link, removing this item would actually decrease the scale's alpha to .694 when considering its theoretical importance to skill loss, while deleting strong items like the fourth and the fifth item would drop the reliability score down to .607 and .572, respectively.

Concerning empirical distortion, the transcripts demonstrate numerous conflicts associated with severe algorithmic errors. For example, P1 reports encountering issues with the reliability of information provided by AI, claiming that "*AI did not always give me reliable information because sometimes it gave me references that looked real, but when I checked them, they were actually inaccurate or invalid*". This problem is echoed by another participant, who reports that AI-based tools often fabricate academic reality, as P4 states that the software "*provided inaccurate information as well as information it made up, and that cost me time to double-check it before including it in my paper*". Speaking about agentic erosion, the transcript analysis reveals instances where students resort to over-reliance on AI owing to excessive academic burden. Specifically, P2 reports that he/she was unable to independently draft papers during particularly busy weeks at university: "*Sometimes I feel burnt out… So sometimes I just throw it on AI to write it for me*". Moreover, P3 notes that it is easy to relinquish cognitive control: "*I realised that AI also understands. Sometimes, it understands better than me, so I let it do it*". Finally, in terms of structural unfairness, participants report concerns with regard to the ethical implications of AI writing and the mechanical nature of language generation. Thus, P1 warns that an unedited submission leads directly to "*plagiarism*", whereas P3 mentions that copying prompts allows one to gain unfair advantages in the process of writing papers, as "*it's not fair because that effort is just to copy a prompt and give it to AI to write. It's not the actual sentence that you use your talent to write*".

4.4. EFL Juniors' Strategies Recommendation to Maintain Academic Integrity

Table 13

Recommended Strategies for Maintaining Academic Integrity

		Responses		Percent of Cases
		N	Percent	
AI strategies to support students[a]	Clear of AI usage guidelines	109	18.4%	61.9%
	Examples of acceptable AI use	110	18.5%	62.5%
	AI citation instructions	99	16.7%	56.3%
	Ethical discussions implementation	105	17.7%	59.7%
	Teachers feedback	107	18.0%	60.8%
	AI ethics training	64	10.8%	36.4%
Total		594	100.0%	337.5%

From Table 13, it is evident that students prioritise pragmatic institutional and mentor guidance rather than training workshops. The most popular approach is to provide "*Examples of acceptable AI use*" (62.5% of the participants). Another top-rated strategy is to introduce "*Clear of AI usage guidelines*" (61.9%). Moreover, students' top preferences include direct teacher guidance, such as "*Teachers feedback*" (60.8%) and "*Ethical discussions implementation*" (59.7%). Additionally, students are in need of instructions on AI citation (56.3%). Formal structured approaches, such as AI ethics training, are not as popular and are preferred by 36.4% of the students. The provided statistics prove that pragmatic examples, clear guidelines, and direct teacher feedback are required by the students for maintaining their academic integrity.

Table 14

Quality of Students' Recommendations

		EFL Juniors' recommendation				Total
		No Opinion	Apprehensive Perspective	Pragmatic Perspective	Visionary Perspective	
Male	Count	12	2	28	13	55
	% within Gender	21.8%	3.6%	50.9%	23.6%	100.0%
Female	Count	12	7	74	28	121
	% within Gender	9.9%	5.8%	61.2%	23.1%	100.0%
Total	Count	24	9	102	41	176
	% within Gender	13.6%	5.1%	58.0%	23.3%	100.0%

Analysing the details of the percentage distribution in Table 14, the most noticeable trend is the preference of both genders for the "*Pragmatic Perspective*". Both the male and female subgroups demonstrate this preference at the highest percentages, 50.9% and 61.2%, respectively. Secondly, one can see the gender neutrality of the "*Visionary Perspective*" category as it also presents itself at almost equal rates, namely, 23.6% of the male respondents and 23.1% of the female respondents. Only minor differences exist among the other categories; the number of "*No Opinion*" responses among males is higher than the number of responses among females by 11.9%, and the "*Apprehensive Perspective*" is equally insignificant within both gender subgroups at 3.6% and 5.8%.

Regarding in-depth interviews, none of the participants was in favour of an outright ban on technologies in the classroom but advocated instead an educational model that would utilise AI technology as an aid to education under strict supervision. Thus, P2 stated, "*You cannot ban your students from using AI… what we can do is show them the appropriate way to use it*". Similarly, P3 agreed and mentioned that it was essential to understand that it was up to students to apply AI properly and with proper supervision of the teachers, who observed whether it helped or harmed the process. P1 added, "*without the rules, students may misuse AI or become too dependent*". Moreover, the students suggested strict checking rules and oral explanations to ensure their

personal accountability in using this technology. As for P2, he shared his department's policy according to which "*we have a rule that every paper, when checking the plagiarism point and academic similarity point, must be under 3%. This is a way you can ensure academic fairness and integrity*". Lastly, P6 suggested performing an oral verification test by stating, "*be honest about using AI for editing. And so actually, if you can't explain the paper… in two minutes without AI. So it's not your paper… You need to rewrite it again*".

5. Discussion

5.1. EFL Juniors' employment of AI tools

Both the survey and interview findings suggest that the use of AI tools has become a common practice among EFL Juniors, as more than 94% of participants use these solutions several times throughout a semester. Specifically, EFL Juniors consistently employ a number of different applications depending on the stage of writing a paper. High-level cognitive activities are performed using such solutions as ChatGPT (90.3%) and Gemini (81.3%), as both of these conversational models are used by students for topic brainstorming and idea generation. Literature analysis and text refinement are facilitated using source-focused applications like NotebookLM and automated rewriting software such as Quillbot (51.1%) and translation applications (39.8%). Finally, mechanical activities, such as proofreading for grammar and other errors, are aided using proofreading software like Grammarly. The most highly appreciated benefits of such digital scaffolding are related to creative thinking (72.2%) and considerable time-saving during research (64.2%). The combined use of conversational models for concept generation and automated software for surface editing generally aligns with Wood, Bruner, and Ross' (1976) theory of cognitive scaffolding, which suggests that temporary tools enable the performance of complicated activities beyond one's reach. Moreover, the extremely high use of ChatGPT and Gemini in this study corroborates the findings of Fetzer et al. (2025), according to whom generative models have swiftly become indispensable cognitive partners of university students around the globe.

5.2. EFL Juniors' Perceptions of Academic Integrity

According to the results, although the findings suggest that participants were generally aware of the automated help provided by technologies, their inner ethical guidelines are extremely strong. In fact, EFL Juniors view the submission of fully unedited AI content as an extremely unethical practice and would not import raw data to AI to automatically analyse it due to security reasons associated with data breaches. In order to preserve their unique voice and not rely too much on technologies, participants do their best to achieve the right balance by viewing AI tools merely as an assistant and rewriting suggestions made by it instead of blindly copying. Nevertheless, excessive university workload causes significant academic burnout in participants and, sometimes, forces them to step outside their boundaries and use AI-generated text to make things easier. Another problem identified in the course of data analysis pertains to AI hallucinations because students frequently encounter the issue of fabricated academic citations that cannot be considered valid or accurate. This behaviour pattern under the conditions of academic pressure corresponds to the situation fraud concept introduced by Sujeewa et al. (2020), which states that under high pressure (in the case of the study, it relates to academic burnout) and opportunities for unethical actions (in this situation, uncontrolled use of AI), ethics breaks down quite easily. Moreover, participants' concerns related to overdependence and lack of skills

correspond to the prediction made by Selwyn (2014) regarding the de-skilling of learners using automated educational systems.

5.3. EFL Juniors' Strategies Recommendation to Maintain Academic Integrity

Mixed-methods data gathered provide compelling support for a learning approach that involves explicit guidance, practical examples, and verification by humans, not prohibitions that will only prove ineffective through technology. First, in the questionnaire, students mainly request examples of acceptable use of AI (62.5%), guidelines provided by departments (61.9%), and feedback provided by teachers (60.8%). Importantly, however, thematic analysis obtained in interviews offers further insight into why the survey respondents feel that way and provides an explanation of how these approaches may be employed. In particular, interviews conducted with P1 and P3 indicate that vague policies do not work; hence, departments should set clear boundaries between editing and text generation. In order to ensure personal responsibility, participants have recommended spending at least one or two classes on AI education (P2, P5) and using automated checking along with verification by a two-minute oral defence (P6).

6. Conclusion

6.1. Major findings

Based on the findings obtained from the study employing a mixed method approach, it can be concluded that the role of artificial intelligence in facilitating the academic writing process of EFL Juniors in HUFLIT is a vital one, which is deeply rooted in their writing processes. Quantitative data indicate that almost all of EFL Juniors frequently employ different conversational models, including ChatGPT and Gemini, as well as automated editing programs, including QuillBot and Grammarly, in completing their research work. All those technological products function as efficient learning assistants that effectively alleviate the feeling of language anxiety, help save time, and enhance paragraph flow through proposing new vocabulary and more complex sentence structures. Nevertheless, the quick technological advancement poses many threats related to the integrity of students' work. They encounter many technical problems, such as AI hallucinations which invent totally fictional citations, as well as suffer from severe academic burnout that sometimes results in over-dependence on technology. Overall, despite the fact that EFL Juniors use various platforms to enhance linguistic abilities, they keep solid boundaries, not letting technology substitute their thoughts.

6.2. Implications

The results of this study provide essential pedagogical implications that are relevant for the future of higher education in the time of generative AI technologies. Firstly, for the teacher, the normalisation of these tools across society calls for a transition from strict policing to practical classroom guidance. The teacher must be transparent and teach students about AI literacy through specific lectures and exercises showing examples of acceptable help from AI and unacceptable content creation. In addition, interactive oral defence sessions must be introduced to check students' ability to learn independently. Secondly, from the perspective of curriculum design, the results imply the need for a total revamp of syllabi in terms of process rather than product evaluation. Curriculum designers must include AI literacy lessons in academic writing classes and focus on assessment criteria based on critical editing and verification of sources. Finally, for institutions, the fact that restrictive approaches do not work becomes clear. Universities need to

have a strong honour code based on such elements as data privacy rules and clear examples of AI usage both in class and in assignments.

6.3. Limitations

Apart from providing an accurate and comprehensive analysis of the data, however, this study comes with additional limitations that have to be taken into account upon interpretation of its findings. First and foremost, data were gathered exclusively within the framework of one department of one institution, namely the Faculty of Foreign Languages at HUFLIT. Given the fact that this study focuses solely on a limited sample of 176 EFL Juniors in Vietnam, there is a possibility that the observed characteristics in terms of frequency of use, awareness, etc., would differ greatly for students who pursue education in other fields, such as business, law, engineering, etc., as well as students who take EFL courses at other institutions. Secondly, since the research is based solely on survey questions and voluntary in-depth interviews, social desirability bias can be considered an inherent limitation of the current study. Some participants could intentionally provide responses that would indicate less frequent AI use or adherence to rules concerning dishonesty at universities in order not to be penalised academically. Lastly, since this study only adopts the descriptive and cross-sectional approach, it is only possible to have a fleeting picture of student attitudes at one particular moment in time. Thus, we will not be able to make any causal connection between the variables, nor will we be able to observe their dynamic development through time.

6.4. Future research

With respect to extending the empirical conclusions obtained from the present research, three essential methodological recommendations for further research in the field of educational studies are to be mentioned. First, future research should involve longitudinal research to reveal whether a continued use of artificial intelligence tools during four years of college studies helps or hinders the development of an independent proficiency in English grammar and vocabulary, as well as critical thinking. Second, further studies need to carry out cross-university comparisons, employing more varied multi-university samples. Specifically, it would be helpful to identify whether differences in university policy on using AI tools affect students' perceptions of academic integrity to different extents among various academic disciplines. Third, further investigation needs to conduct experiments, comparing the effect of different technological tools on EFL Juniors and their academic integrity. In particular, more research should focus on studying the unique effects of different software ecosystems in influencing the process of students; in particular, the automated literature mapping tools (Connected Papers, Research Rabbit, Litmaps) that help in synthesising texts, as well as statistical packages (SPSS, Stata, and JMP). Future studies may extend the present findings by employing more advanced statistical techniques or approaches. Future studies may develop and validate more comprehensive measurement scales through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to structurally analyse the actual underlying dimensions of digital literacy, ethical boundaries, and automated scaffolding dependencies in language learners.

References

Aggarwal, R., & Ranganathan, P. (2019). Study designs: Part 2 – Descriptive studies. *Perspectives in Clinical Research*, *10*, 34-36. 10.4103/picr.PICR_154_18

Aljuaid, H. (2024). The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review. *Arab World English Journal (AWEJ)*, 26-55. https://dx.doi.org/10.24093/awej/ChatGPT.2

Barrot, J. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745

Bjelobaba, S., Waddington, L., Perkins, M., Foltýnek, T., Bhattacharyya, S., & Weber-Wulff, D. (2024). Research Integrity and GenAI: A Systematic Analysis of Ethical Challenges across Research Phases, *arXiv preprint*. https://doi.org/10.1007/s40979-025-00191-w

Chan, C. K. Y. (2024). Students' perceptions of 'AI-giarism': investigating changes in understandings of academic misconduct. *Education and Information Technologies*, *30*(6), 8087-8108. https://doi.org/10.1007/s10639-024-13151-7

Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Fetzer, D., Gimpel, H., Meindl, O., & Strickmann, J. (2025). Responsible Engineering of Information Systems Based on Generative Artificial Intelligence: An Action Design Research Study at a German Premium Car Manufacturer. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-025-00950-6

Fowler, D. (2023). AI in Higher Education: Academic Integrity, Harmony of Insights, and Recommendations. *Journal of Ethics in Higher Education*, (3), 127–143. https://doi.org/10.26034/fr.jehe.2023.4657

Gayed, J. M. (2025). Educators' perspective on artificial intelligence: equity, preparedness, and development. *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2024.2447169

Hernandez, D., Pasha, L., Yusuf, D., Nurfaizi, R., & Julianingsih, D. (2024). The Role of Artificial Intelligence in Sustainable Agriculture and Waste Management: Towards a Green Future. *International Transactions on Artificial Intelligence (ITALIC)*, *2*(2), 150-157. https://doi.org/10.33050/italic.v2i2.552

Hossain, Z., Çelik, Ö., & Hınız, G. (2025). Exploring EFL Students' AI Literacy in Academic Writing: Insights into Familiarity, Knowledge and Ethical Perceptions. *Kuramsal Eğitimbilim*, *18*(1), 157-181. https://doi.org/10.30831/akukeg.1538011

Issembayeva, N., Ramazanova, D., Unerbayeva, Z., Kuanysheva, Z., & Seiitkazy, P. (2026). Digital etiquette and belief-oriented technology use among Kazakhstani pre-service teachers: a cross-sectional survey. *Frontiers Education*, *11*. https://doi.org/10.3389/feduc.2026.1827729

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-97652-6

Khalifa, M., & Albadaway, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, *5*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

McCabe, D. (2005). Cheating among college and university students: A North American perspective. *International Journal for Educational Integrity*, *1*(1). https://doi.org/10.21913/IJEI.v1i1.14

Nordlof, J. (2014). Vygotsky, Scaffolding, and the Role of Theory in Writing Center Work. *Writing Center Journal*, *34*(1). https://doi.org/10.7771/2832-9414.1785

Nyaaba, M., Wright, A., & Choi, G. (2025). Generative AI and Power Imbalances in Global Education: Frameworks for Bias Mitigation. *Cornell University.*

Richards, D., Worden, D., Song, X., & Lavorel, S. (2024). Harnessing generative artificial intelligence to support nature-based solutions. *People and Nature*, *6*, 882-893. https://doi.org/10.1002/pan3.10622

Sujeewa, G., & Kawshalya, M. (2020). Detecting red flags of corporate financial statement frauds using the Beneish M score model in Sri Lanka. *International Journal of Accounting and Business Finance*, *6*(2), 71-81. https://doi.org/10.4038/ijabf.v6i2.72

Torres, P. J., & Kahveci, Y. E. (2025). Effectiveness of Artificial Intelligence (AI) in language teaching. *Computers and Education: Artificial Intelligence*, *9,* 100522. https://doi.org/10.1016/j.caeai.2025.100522

Tuan, N. (2026). Boosting Fluency, Reducing Fear: Utilizing Generative AI as a Scaffolding Tool in EFL Speaking Classes. *MSI Journal of Education and Social Science*, *2*(1), 01–13. https://doi.org/10.5281/zenodo.18448552

Bui, T. T. U., & Tong, T. V. A. (2025). The Impact of AI Writing Tools on Academic Integrity: Unveiling English-Majored Students' Perceptions and Practical Solutions. *AsiaCALL Online Journal, 16*(1), 83-110. https://doi.org/10.54855/acoj.251615

Wood, D., Bruner, J., & Ross, G. (1976). The Role Of Tutoring In Problem Solving. *The Journal of Child Psychology and Psychiatry*, *17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Biodata

Nguyen Le Anh Tuan is an accomplished alumnus of Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), having successfully graduated from the Faculty of Foreign Languages, majoring in Translation and Interpretation. Driven by a profound academic curiosity, he possesses a strong passion for conducting empirical research and writing scientific papers centered on modern educational technology.

Thai Thi Ngoc Huyen is a dedicated alumnus of Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). She graduated from the Faculty of Foreign Languages, majoring in Pedagogy. Deeply passionate about the field of education, she has

committed herself to developing effective teaching methods that inspire and motivate language learners.

Ms. Diep Minh Chau holds a Master's degree in TESOL from the University of Canberra, Australia. She is currently a permanent lecturer at HUFLIT, specializing in English Language Teaching (ELT), and has also taught at the Vietnam USA Society English Centers (VUS) for over sixteen years. She is deeply committed to teacher training for English Language Teaching and English for Young Adults programs. Additionally, she serves as an academic advisor for pedagogy students at HUFLIT, guiding them through their teaching internships at high schools.

#113 CHALLENGES AND PEDAGOGICAL RESPONSES IN TRANSITIONING TO ESL-ORIENTED TEACHING: A MIXED-METHODS CLASSROOM CASE STUDY IN A VIETNAMESE UNIVERSITY

Huỳnh Thị Thùy Dung[1]

Abstract

The transition from English as a Foreign Language (EFL) to ESL-oriented teaching has become a key direction in Vietnamese higher education, driven by national policies emphasizing communicative competence. However, implementing this shift remains challenging in contexts where traditional, form-focused instruction persists. This study examines ESL-oriented teaching in a Vietnamese university through a mixed-methods case study conducted in two General English classes (n = 82). Guided by three research questions, the study investigates students' perceptions of ESL-oriented teaching, teacher adaptation during the EFL–ESL transition, and the challenges and opportunities emerging in classroom practice. Data were collected over 12 weeks using questionnaires, classroom observations, and teacher reflective journals. Findings show that students held moderately positive perceptions of ESL-oriented teaching, particularly regarding its usefulness for communication. However, affective barriers such as low confidence, fear of mistakes, and reluctance to participate limited engagement. Teacher adaptation was mainly reflected in scaffolding and task-based instructional redesign. Despite initial difficulties, student engagement and speaking confidence gradually improved. The study concludes that the transition to ESL-oriented teaching is gradual and context-dependent, highlighting the central role of teacher agency in mediating pedagogical change and the need for sustained scaffolding and supportive classroom environments.

Keywords: ESL transition, teacher agency, classroom challenges, student engagement, communicative teaching, case study

1. Introduction

1.1. Background

Across many EFL contexts, developing learners' communicative competence is still a challenge, even after long-term curriculum reforms. In many classrooms, teaching is still focused on grammar and exams, which creates a gap between policy and real practice. Communicative Language Teaching (CLT) and Task-Based Language Teaching (TBLT) are widely used to improve interaction and meaningful use of language (Ellis, 2003; Long, 2015). However, how they work in class often depends more on context than on theory alone (Burns & Richards, 2021).

Recent studies therefore focus more on teachers as key actors in reform. Teachers do not only follow methods but also adjust them based on their classroom situation. This is often explained through teacher agency, which refers to teachers' ability to make decisions in real

teaching contexts (Priestley et al., 2015; Tao & Gao, 2021). In Vietnam, English reform under the National Foreign Languages Project (2025-2035, vision 2045) promotes communicative competence. However, in higher education, implementation is still uneven because of many classroom constraints that affect teachers' practices (Dang et al., 2024; Ngo, 2024).

1.2. Problem Statement and Research Gap

Despite strong policy support for communicative teaching, the shift from traditional EFL to more ESL-oriented instruction in Vietnamese universities is still difficult. Previous studies point to common problems such as exam pressure, limited classroom interaction, and institutional constraints (Nguyen et al., 2023). More recent work suggests that teacher agency plays an important role in how reforms are implemented in real classrooms (Tao & Gao, 2021; Ngo, 2024). However, there is still limited classroom-based evidence, especially studies using multiple data sources to capture what actually happens in teaching practice.

Although communicative language teaching and teacher agency have been studied separately, less is known about how teacher agency shapes the transition from EFL to ESL-oriented instruction in real classroom contexts in Vietnamese higher education. This study addresses this gap by examining both teacher adaptation and student responses through a mixed-methods classroom case study.

1.3. Research Objectives

This study examines ESL-oriented teaching in a Vietnamese university classroom. It aims to:

- explore students' views on ESL-oriented teaching;
- look at how the teacher adapted from EFL to ESL-oriented instruction;
- identify main challenges and opportunities in using communicative teaching.

1.4. Delimitations

This study was carried out in one university classroom over a semester. It focuses on one teacher and one student group, using classroom observation, teacher reflection, questionnaires, and informal feedback. The results are specific to this context and are not meant to be generalised widely.

1.5. Research Questions

RQ1. How do students perceive ESL-oriented teaching in a Vietnamese university classroom?

RQ2. How does the teacher adapt instruction during the transition from EFL to ESL-oriented teaching?

RQ3. What challenges and opportunities emerge during this implementation?

2. Literature Review

2.1. The Shift From EFL to ESL-Oriented Teaching

In recent years, the difference between English as a Foreign Language (EFL) and English as a Second Language (ESL) has gained more attention in language teaching. EFL usually focuses

on learning English for exams and classroom use, while ESL-oriented teaching emphasizes real communication and interaction. This change is linked to globalization and the increasing use of English in international settings.

In Vietnam, the move toward ESL-oriented teaching has been driven by national goals related to global integration, employability, and educational reform. As a result, English teaching has gradually shifted from grammar-based instruction toward more communicative and learner-centered approaches. However, this transition remains uneven across institutions and classrooms, as it requires not only curriculum change but also substantial pedagogical adjustment at the classroom level. This macro-level shift provides the policy background of the present study, which assumes that curriculum reform is not directly implemented but mediated through classroom actors, particularly teachers.

2.2. Pedagogical Implications of ESL-Oriented Teaching

The shift toward ESL-oriented teaching requires changes in how classrooms are run. Communicative Language Teaching (CLT) and Task-Based Language Teaching (TBLT) are often linked to this approach because they focus on real communication and meaningful use of language. Previous studies show that these methods can improve students' interaction and engagement. Ellis (2003) notes that task-based teaching helps develop communicative competence through real communication, while Nunan (2004) stresses the need for tasks that match real-life use of English. However, most of these studies focus more on teaching principles than on how teachers actually adapt them in real classrooms.

Willis (1996) further proposes a structured framework for task implementation, including pre-task preparation, task cycle, and post-task language focus. Such frameworks illustrate how communicative pedagogy can be operationalised in classroom settings. However, as classroom realities vary, these pedagogical models are often adapted rather than directly applied. This adaptation process is central to the present study's conceptual framework, which views pedagogical practice as being shaped by teacher interpretation and contextual constraints.

2.3. Challenges and Opportunities in Transitioning to ESL-Oriented Teaching

Despite the benefits of ESL-oriented teaching, previous studies still report many challenges, especially in EFL contexts. Common issues include exam pressure, rigid curricula, and limited space for communicative activities (Vu & Nguyen, 2021). Teacher-related factors such as lack of training, limited resources, and difficulty managing interactive tasks also make implementation harder (Lam et al., 2021). In addition, students' past experience with teacher-centered learning can lead to low confidence, speaking anxiety, and low participation.

However, although previous studies have identified both the challenges and the potential benefits of ESL-oriented teaching, relatively little attention has been paid to how university teachers experience the transition in specific institutional contexts. In particular, limited research has examined how lecturers navigate the shift to ESL-oriented teaching, the challenges they encounter during implementation, and the practical strategies they employ to address these challenges in Vietnamese higher education. Addressing this gap is important for understanding how pedagogical reforms can be effectively implemented in local educational settings.

2.4. Conceptual Framework and Research Gap

Although previous studies have pointed out many challenges in ESL-oriented teaching, there are still few classroom-based studies showing how these challenges are handled in real teaching practice. In particular, not much attention has been given to how teachers understand, adjust, and respond to pedagogical change in their actual classrooms.

Figure 1

Conceptual Framework of ESL-Oriented Teaching Implementation in a Vietnamese University Classroom

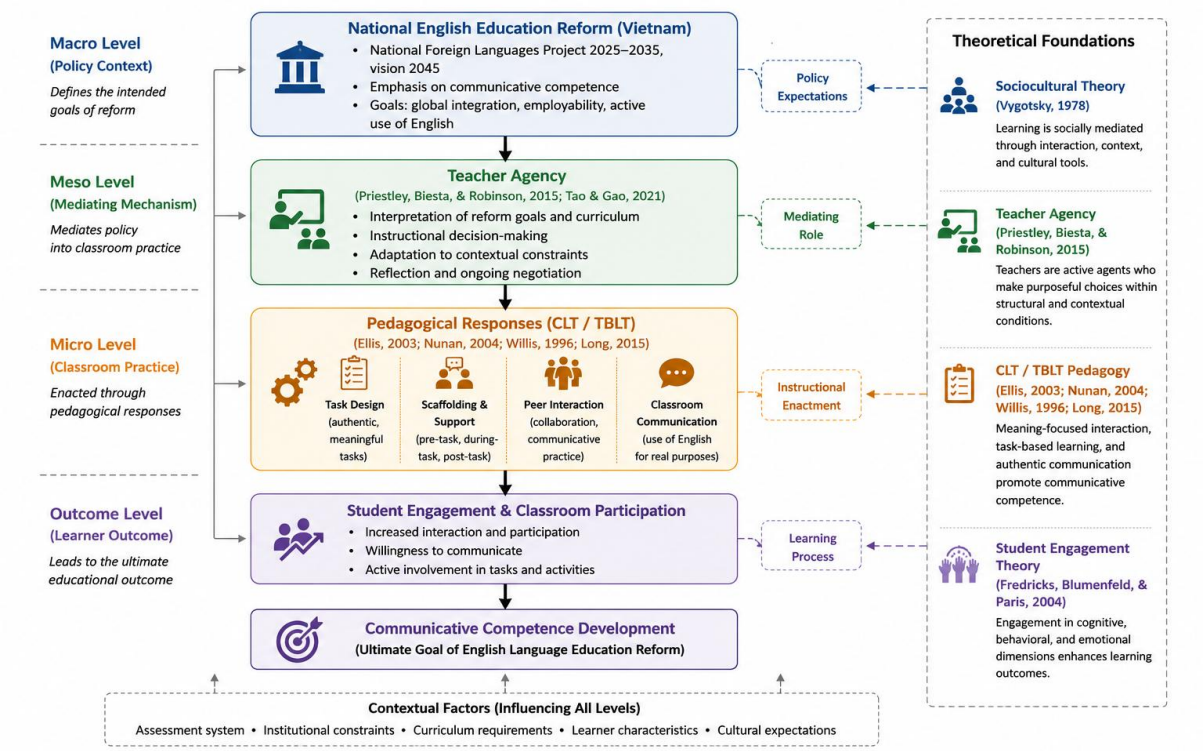


Figure 1. Conceptual framework of ESL-oriented teaching implementation in a Vietnamese university classroom.

Grounded in this gap, the present study adopts a process-oriented conceptual framework. At the macro level, national English education reform in Vietnam promotes communicative competence as an intended outcome of policy. However, the implementation of this goal is not direct but mediated through classroom practice.

At the meso level, teacher agency plays a central role in this mediation process. Teachers interpret reform expectations and make pedagogical decisions based on contextual constraints. These decisions are reflected in their adaptation of CLT and TBLT through task design, scaffolding strategies, and classroom interaction patterns (Priestley, Biesta, & Robinson, 2015; Tao & Gao, 2021).

At the micro level, these pedagogical responses influence student engagement and participation, which in turn contribute to communicative competence development (Richards, 2006). This suggests a dynamic relationship between policy, teacher practice, and learner outcomes.

Therefore, the three dimensions of this framework directly inform the research questions of the study. Student perceptions of ESL-oriented teaching reflect the micro-level learner experience. Teacher adaptation during instructional transition represents meso-level teacher agency in practice. Finally, the challenges and opportunities emerging from classroom implementation illustrate the interaction between contextual constraints and pedagogical possibilities. Together, these dimensions provide an integrated lens for examining ESL-oriented teaching in a Vietnamese university context.

3. Methodology

3.1. Research Design

This study used a mixed-methods case study design, combining both qualitative and quantitative data to better understand the shift from EFL to ESL-oriented teaching. The case study was chosen to explore classroom practice, teacher adaptation, and student responses in one specific context.

In addition, a reflective practice component was incorporated, as the teacher-researcher continuously evaluated and adjusted instructional strategies throughout the implementation process. This design aligns with the study's aim to examine pedagogical change as a situated and dynamic process.

3.2. Research Context and Participants

3.2.1. Context

The study was conducted at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), Vietnam. The institution provides General English courses for non-English major students as part of its foundational curriculum.

3.2.2. Participants

The participants were 82 first-year non-English major students from two General English classes. They were 18-19 years old and had mixed English levels, from elementary to pre-intermediate. Both classes used the same curriculum and were taught by the same teacher to keep instruction consistent. Most students had experienced teacher-centered, grammar-focused learning in high school, with little practice in real communication. Convenience sampling was used because both classes were taught by the researcher during the semester.

3.2.3. Participant demographics

Table 1

Participant Demographics

Variable	**Category / Mean**
Gender	Male: 36 (43.9%) / Female: 46 (56.1%)
Age	18-19 years old
CEFR level	A1-A2 (majority), some A2+
Major	Non-English majors (Business, IT, Tourism, etc.)
Years learning English	7-10 years (mean ≈ 8.5 years)

3.3. Instructional Procedures

The shift to ESL-oriented teaching was carried out over about 12 weeks. During this period, traditional teaching was gradually changed to include more communicative and task-based activities. Lessons were redesigned to focus on real-life communication and student interaction, such as role-plays, pair and group discussions, problem-solving tasks, and speaking activities linked to students' daily situations. The teaching followed a task-based structure (Willis, 1996), including pre-task, task, and post-task stages. To support the change, the teacher used simple strategies like scaffolding language, giving clear instructions, organizing group work, and creating a supportive classroom atmosphere to reduce anxiety and encourage participation.

3.4. Data Collection

Data were collected from three sources to ensure triangulation:

3.4.1. Classroom observations

Used to document student engagement, interaction patterns, and participation during ESL-oriented tasks.

3.4.2. Teacher reflective journals

Used to record instructional decisions, classroom challenges, and ongoing pedagogical adjustments.

3.4.3. Student questionnaire

A structured questionnaire was administered to 82 students after the intervention.

- **Total items:** 15 items
- **Scale:** 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree)
- **Sections:**
  - Student engagement (5 items) (Fredricks et al., 2004)
  - Speaking confidence (5 items) (Woodrow, 2006)
  - Perceptions of ESL-oriented teaching (5 items) (Richards, 2006)

The questionnaire was adapted from previous validated instruments in studies on communicative language teaching and learner engagement (e.g., Fredricks et al, 2004; Woodrow, 2006; Richards, 2006). A pilot study with 10 students (not included in the main sample) was conducted to ensure clarity and appropriateness of items. Minor wording adjustments were made based on feedback.

3.5. Validity and Reliability

3.5.1. Validity

Content validity was ensured through expert review by two specialists in applied linguistics and TESOL. They evaluated the questionnaire in terms of clarity, relevance, and alignment with research objectives. Minor revisions were made based on their feedback. A pilot test was also conducted to identify ambiguous wording and improve item clarity.

3.5.2. Reliability

Internal consistency was tested using Cronbach's Alpha. All values were above .80, showing good reliability (Hair et al., 2019). The overall questionnaire had a reliability of $\alpha = .87$, which is considered high. The subscale results were as follows:

- Engagement: $\alpha = .84$
- Speaking confidence: $\alpha = .86$
- Perception of ESL teaching: $\alpha = .85$

3.6. Data Analysis

Quantitative and qualitative data were analyzed separately and then integrated for interpretation.

3.6.1. Quantitative analysis

Quantitative data from the questionnaire were analyzed using IBM SPSS Version 26. Descriptive statistics (means and standard deviations) were used to describe students' responses. Paired-samples t-tests were then used to compare students' engagement, speaking confidence, and willingness to participate before and after the intervention. The significance level was set at $p < .05$. When normality assumptions were not met, the Wilcoxon signed-rank test was used instead.

3.6.2. Qualitative analysis

Qualitative data from classroom observations, reflective journals, and open-ended questionnaire responses were analyzed using thematic analysis following a structured coding procedure:

3.6.2.1. Open coding

Initial coding of raw data to identify meaningful units related to teaching practices, student responses, and classroom interaction.

3.6.2.2. Axial coding

Grouping of codes into categories based on relationships and patterns (e.g., engagement, anxiety, scaffolding strategies).

3.6.2.3. Theme development

Emergence of broader themes such as:

- Student engagement transformation
- Teacher adaptation strategies
- Classroom interaction dynamics
- Implementation challenges

3.6.2.4. Cross-checking/ triangulation

Themes were cross-validated across data sources (observations, reflections, questionnaires) to ensure consistency.

3.6. *Ethical Considerations*

Participation in the study was voluntary. Students were informed of the purpose of the study, and all questionnaire responses were collected anonymously. The data were used solely for research purposes, and participants' confidentiality was maintained throughout the study.

4. Results

4.1. Overview of Findings

The findings combine both quantitative data (questionnaires) and qualitative data (classroom observations, teacher reflection, and student feedback) to answer the three research questions. They are organized by research question. First, the quantitative results are presented to show general trends in students' responses, followed by qualitative findings to explain these results in more detail. Using both types of data helps give a clearer and more complete picture of the shift toward ESL-oriented teaching.

4.2. Students' Perceptions of ESL-Oriented Teaching (RQ1)

This section presents findings related to students' perceptions of ESL-oriented teaching based on questionnaire data and supporting qualitative evidence.

4.2.1. Quantitative results

Table 2 presents descriptive statistics of students' perceptions.

Table 2

Item-Level Results of Students' Perceptions of ESL-Oriented Teaching (n = 82)

Item	Mean	SD
I enjoyed participating in communicative tasks.	3.84	0.73
ESL-oriented activities increased my classroom engagement.	3.81	0.69
I was willing to communicate with classmates in English.	3.55	0.71
I felt more confident speaking English.	3.39	0.77
The communicative tasks were useful for real-life communication.	3.96	0.61

Among the five questionnaire items, students rated the usefulness of communicative tasks highest (M = 3.96, SD = 0.61), showing that they saw clear practical value. In contrast, speaking confidence was the lowest (M = 3.39, SD = 0.77), suggesting that emotional barriers still existed. This means students liked communicative activities, but they were still not fully confident when speaking English. Overall, students had positive views of the ESL-oriented approach, but still felt some hesitation in oral communication.

4.2.2. Qualitative results

Theme 1: Initial anxiety and low confidence

Codes:

- nervousness when speaking
- fear of mistakes
- avoidance of speaking tasks

Representative quotes:

"I feel nervous when I have to speak English in front of others." (Student feedback)

"Many students remained silent during early speaking activities." (Observation, Week 2)

Qualitative data from classroom observations and student feedback showed that anxiety and fear of making mistakes were the main problems at the beginning of the implementation. This helps explain the low score for speaking confidence (M = 3.41). At the same time, students still viewed ESL-oriented teaching positively, especially in terms of task usefulness and classroom engagement. This shows that positive attitudes existed, but were still affected by emotional barriers. Overall, both quantitative and qualitative results consistently indicate that students were generally positive about ESL-oriented teaching, even though affective challenges remained.

4.3. Teacher Adaptation During the Transition (RQ2)

This section reports how the teacher adapted instruction from EFL to ESL-oriented teaching based on qualitative data.

Theme 2: Scaffolding as a key adaptive strategy

Codes:

- providing model sentences
- vocabulary support
- step-by-step instructions
- guided practice

Representative quotes:

"After providing sample expressions, students participated more actively." (Observation, Week 6)

"Clear scaffolding helped students understand the task better." (Reflective journal)

Teacher scaffolding played a crucial role in supporting students' transition to communicative learning. It reduced cognitive difficulty and increased participation.

Theme 3: Task-based restructuring of classroom activities

Codes:

- role-play tasks
- group discussions
- real-life communication tasks
- peer interaction

Representative quotes:

"Students worked in groups to solve real-life communication tasks." (Observation, Week 7)

The teacher gradually moved from form-focused teaching to task-based learning, following ESL-oriented principles. This change was mainly seen in how the teacher used scaffolding and redesigned classroom activities to be more communicative and task-based. Overall, the shift happened step by step, not suddenly. Scaffolding and task redesign were the two main ways the teacher adapted instruction.

4.4. Challenges and Opportunities in Implementation (RQ3)

4.4.1. Qualitative evidence

This section presents both difficulties and positive outcomes during the transition.

Theme 4: Affective and behavioral challenges

Codes:

- fear of mistakes
- low participation
- dependence on teacher
- silence in group work

Representative quotes:

"Some students waited for the teacher instead of continuing the task independently." (Observation, Week 4)

"I am afraid of speaking incorrectly." (Student feedback)

Affective barriers and prior learning habits limited student participation, especially in the early stages of implementation.

Theme 5: Uneven participation in group work

Codes:

- dominant students
- passive group members
- language switching to Vietnamese

Representative quotes:

"Only one or two students actively led the discussion in each group." (Observation, Week 5)

Participation inequality was observed in group tasks, particularly in larger classes where teacher monitoring was limited.

Theme 6: Pedagogical opportunities and improvement

Codes:

- increased engagement over time
- more voluntary participation

- improved interaction
- growing confidence

Representative quotes:

"More students started participating in discussions in later weeks." (Observation, Week 8)

"I feel more confident speaking with my group now." (Student feedback)

Despite initial challenges, ESL-oriented teaching gradually improved student engagement and interaction. Structured tasks and scaffolding contributed to increased willingness to communicate.

4.4.2. Quantitative evidence of improvement

Table 3

Pre-Post Comparison of Student Engagement

Variable	Pre Mean (SD)	Post Mean (SD)	t	p	Cohen's d
Willingness to speak	2.95 (.70)	3.78 (.72)	6.42	<.001	0.71
Active participation	3.01 (.68)	3.85 (.70)	6.87	<.001	0.76
Speaking confidence	2.88 (.72)	3.62 (.74)	5.98	<.001	0.66

Paired-samples t-tests revealed statistically significant improvements across all three indicators.

- Willingness to speak: $t(81) = 6.42$, $p < .001$
- Participation: $t(81) = 6.87$, $p < .001$
- Speaking confidence: $t(81) = 5.98$, $p < .001$

The biggest improvement was in active participation ($d = 0.76$), followed by willingness to speak ($d = 0.71$) and speaking confidence ($d = 0.66$). According to Cohen (1988), these are moderate to large effects, showing that the ESL-oriented teaching had a positive impact on students' communication.

The transition involved both challenges (affective barriers, uneven participation) and opportunities (increased engagement and improved communicative behavior over time). Taken together, evidence from the present study suggests that although the transition created considerable challenges, appropriate pedagogical support enabled measurable improvements in classroom engagement and willingness to communicate.

4.5. Overall Summary of Findings

The results show that ESL-oriented teaching was perceived positively by students but initially constrained by affective factors such as anxiety and low confidence. Teacher adaptation, particularly through scaffolding and task-based redesign, played a central role in facilitating classroom participation. Although challenges persisted, quantitative and qualitative evidence both indicate a gradual improvement in engagement and communicative interaction.

5. Discussion

5.1. Transition as a Gradual and Non-Linear Process

The findings show that the shift from EFL to ESL-oriented teaching happens slowly, not right away. At the start, many students were not very willing to join communicative tasks because they were used to teacher-centered and grammar-focused learning. This suggests that changing activities alone is not enough to change how students behave in class. As Richards (2006) notes, communicative teaching requires students to take a more active role in using English for real communication. Burns and Richards (2021) also point out that how CLT is applied depends a lot on classroom context, not just teaching methods. Therefore, students' initial hesitation should be viewed as a natural stage of adaptation rather than a sign of unsuccessful implementation. This interpretation is consistent with previous studies showing that communicative reforms require sustained pedagogical and learner adjustment rather than immediate behavioral change (Lam et al., 2021). However, unlike earlier research, which has mainly examined reform at the policy or institutional level, the present study provides classroom-level evidence of how teacher adaptation and learner participation evolved during a 12-week implementation.

From the perspective of the conceptual framework, this gradual transition primarily reflects micro-level adaptation, where learners progressively adjusted to communicative classroom practices rather than responding immediately to pedagogical change. At the same time, the findings suggest that these learner responses were shaped by meso-level teacher mediation through continuous instructional adjustment, illustrating the dynamic relationship proposed in Figure 1.

5.2. Affective Barriers and the Legacy of EFL Instruction

Affective factors, especially anxiety and fear of making mistakes, were the main barriers to student participation. They affected how willing students were to speak in different classroom tasks. This shows that communicative ability is not only about language skills but also about feelings and confidence. In learning environments that focus strongly on correctness, students often see speaking English as risky rather than helpful. Previous studies also report speaking anxiety as a common problem in communicative classrooms (Woodrow, 2006; Vu & Nguyen, 2021). The present study confirms these findings but further demonstrates that such affective barriers gradually diminished when communicative tasks were supported by scaffolding and a low-anxiety classroom environment. This highlights the importance of considering affective support as an integral component of ESL-oriented pedagogy rather than a supplementary classroom practice.

These findings further reinforce the micro-level dimension of the conceptual framework by demonstrating that learners' communicative participation is influenced not only by linguistic competence but also by affective conditions created through classroom interaction.

5.3. Learner Dependency and Challenges to Autonomy

Another issue from the findings is that students still depended a lot on teacher guidance. Even during task-based activities, many students expected clear instructions, sample answers, and quick correction from the teacher. This matches Willis (1996), who notes that students new to TBLT often find it hard when they have less teacher control and more responsibility for learning. The results also show that students in this study had not yet fully developed the autonomy needed for ESL-oriented learning. Similar findings are reported in Vietnamese higher education, where students still rely heavily on teachers even in communicative classes (Ngan et al., 2025). Unlike previous studies that mainly describe learner dependency as a contextual challenge, this study

illustrates how structured scaffolding can gradually encourage greater learner autonomy during classroom interaction.

Within the conceptual framework, learner dependency represents a micro-level challenge that requires sustained meso-level pedagogical support. Teacher scaffolding functioned as the mediating mechanism through which greater learner autonomy gradually emerged.

5.4. Contextual Constraints: Large Classes and Interactional Inequality

The study also underscores the impact of contextual constraints, particularly large class size and mixed proficiency levels. Uneven participation, with more proficient students dominating discussions, was a recurring issue. This finding is consistent with Ellis (2003), who acknowledges that managing interaction in large classes poses significant challenges for TBLT implementation. Similarly, Nunan (2004) notes that class size can limit opportunities for meaningful interaction and individualized feedback. In Vietnam, Nguyen et al. (2023) also identifies large-class teaching as a major barrier to implementing student-centered approaches. The present study adds nuance to this perspective by showing that, while large classes constrain interaction, they do not necessarily prevent the development of communicative competence if appropriate strategies are employed.

These contextual constraints illustrate how macro-level educational conditions, such as institutional realities associated with the implementation of national language policy, continue to shape classroom practice. However, consistent with the proposed framework, their influence was mediated through teachers' instructional decisions rather than determining learning outcomes directly.

5.5. Teacher Agency and the Role of Reflective Practice

A key contribution of this study is the role of teacher agency in the transition process. Through ongoing reflection and adjustment, the teacher could respond to classroom challenges and change teaching strategies when needed. This supports the idea of reflective teaching, where teachers are seen as active decision-makers, not just people who follow policy. As Nunan (2004) notes, teaching innovation depends on how teachers adapt theory to fit their own classrooms. This idea is also supported by Tao and Gao (2021) and Ngo (2024), who argue that reform depends on teachers' professional judgment in practice. Unlike earlier studies that focus mainly on theory, this study provides classroom evidence showing how reflection, scaffolding, and small teaching adjustments helped support the shift toward ESL-oriented teaching.

Among the three dimensions of the conceptual framework, the findings provide the strongest support for the central mediating role of meso-level teacher agency. Rather than functioning as passive recipients of policy, teachers actively interpreted reform goals, adapted instructional practices, and responded to emerging classroom needs. This mediating role explains how macro-level policy aspirations were translated into observable micro-level learner engagement.

5.6. The Effectiveness of Scaffolding and Real-World Tasks

Despite initial challenges, observed classroom practices reveal that scaffolding, structured interaction, and authentic communicative tasks were effective in promoting student engagement and willingness to communicate. This supports Willis's (1996) task-based framework, which highlights the role of pre-task support in facilitating successful task performance, and is consistent

with Ellis (2003) and Nunan (2004), who emphasize the value of authentic language use. Similarly, Vu and Nguyen (2021) found that Vietnamese learners were more motivated when classroom activities reflected real-life communication. The present study extends these findings by providing both qualitative evidence and significant paired-samples t-test results showing that these pedagogical strategies contributed to measurable improvements in students' classroom participation and speaking confidence over time.

Viewed through the conceptual framework, scaffolding represents the instructional mechanism through which teacher agency connects policy intentions with learner participation, thereby linking the meso and micro levels of implementation.

5.7. Implications for ESL-Oriented Teaching in Vietnam

The findings have implications across all three levels of the conceptual framework. At the macro level, national ESL policy should be accompanied by curriculum alignment, assessment reform, and sustained institutional support to facilitate meaningful classroom implementation. At the meso level, teacher professional development should strengthen teachers' capacity to exercise agency through reflective practice, scaffolding, and adaptive pedagogical decision-making. At the micro level, classroom instruction should prioritize learner engagement, affective support, and authentic communicative opportunities to promote students' willingness to communicate and communicative competence. Taken together, these findings support the conceptual framework by demonstrating that successful ESL-oriented teaching depends on the dynamic interaction among policy intentions, teacher agency, and learner participation rather than on any single dimension in isolation.

6. Conclusion

6.1. Summary of Key Findings

This study investigated the transition from EFL-oriented to ESL-oriented teaching in a Vietnamese university context using a mixed-methods case study design. The findings reveal that this transition is complex, gradual, and strongly shaped by both learner-related affective factors and contextual constraints.

At the student level (RQ1), many learners were not confident at first. They were afraid of making mistakes and not very willing to speak in communicative tasks. However, questionnaire results still showed that students had fairly positive views of ESL-oriented teaching, especially its usefulness.

At the teacher level (RQ2), change mainly came through scaffolding and redesigning tasks to be more communicative. These adjustments helped connect traditional teaching with ESL-oriented practices.

At the implementation level (RQ3), the study found both difficulties and benefits. Problems included anxiety, dependence on the teacher, and unequal participation in large classes. At the same time, both quantitative and qualitative data showed that students gradually became more engaged and more confident in speaking over the 12-week period.

6.2. Theoretical Contributions

This study contributes to ESL-oriented teaching research in three ways. First, it extends TBLT and CLT theories (Ellis, 2003; Nunan, 2004; Willis, 1996) by showing that tasks and interaction alone are not enough. Students still need clear scaffolding that matches their level and makes tasks easier to follow. Second, it shows that student participation depends not only on language ability but also on feelings like anxiety and confidence. This means affective factors should be included in models of communicative teaching. Third, it highlights teacher agency in real practice. Teachers do not just follow policy, but adjust and adapt it to fit their classroom context in Vietnamese universities.

6.3. Pedagogical and Policy Implications

Several implications can be drawn from the findings. First, the move to ESL-oriented teaching should be seen as a slow process. It needs time and support for both teachers and students. Quick changes are unlikely to work well if students' habits and classroom culture do not change at the same time. Second, affective factors like anxiety and fear of making mistakes need to be addressed directly. A safe and supportive classroom environment is important to help students feel more willing to speak. Third, scaffolding and clear task design are very important, especially in large classes where students may find it hard to work independently. Finally, teacher reflection and agency are key in making this change work. Teachers need space to adjust and adapt teaching in real classroom conditions. From a policy view, ESL reform should go beyond changing the curriculum. It should also include ongoing teacher training, practical classroom support, and assessment that values communication. Institutional support for reflective teaching is also necessary to help teachers apply reforms effectively in different contexts.

6.4. Limitations of the Study

Despite its contributions, this study has several limitations. First, it was conducted with a relatively small sample of students from a single institution, which may limit the generalizability of the findings. Second, the study relied partly on self-reported data, which may be subject to bias. Third, the duration of the implementation was relatively short, and therefore the long-term impact of ESL-oriented teaching practices was not fully explored.

6.5. Recommendations for Future Research

Future research could expand this study by looking at ESL-oriented teaching in different universities and teaching contexts. Long-term studies are also needed to see how students' speaking skills and learning behaviors change over time. In addition, future work could explore how digital tools and AI can support ESL-oriented teaching, especially in large classes.

6.6. Final Remarks

In conclusion, this study shows that the shift from EFL to ESL-oriented teaching is both difficult and changeful. At first, students face problems such as anxiety, low confidence, and dependence on the teacher. However, these issues can improve over time with scaffolding, task-based activities, and reflective teaching.

The findings suggest that successful change does not depend only on new teaching methods, but also on how teachers act, how students adapt, and the classroom context. ESL-oriented teaching should therefore be seen not as a final goal, but as an ongoing process of change in classroom practice.

References

Burns, A., & Richards, J. C. (2021). *The Cambridge guide to learning English as a second language*. Cambridge University Press. https://doi.org/10.1017/9781009024761

Dang, T. K. A., Bohlinger, S., & Cong-Lem, N. (Accepted/In press). Teachers' transformative agency in English-medium instruction in higher education in Vietnam: a cultural-historical theory perspective. *Journal of Multilingual and Multicultural Development*. https://doi.org/10.1080/01434632.2024.2349803

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). *School engagement: Potential of the concept, state of the evidence*. Review of Educational Research, 74(1), 59-109. https://doi.org/10.3102/00346543074001059

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Lam, H. T. L., Nguyen, S. V., & Nguyen, H. A. T. (2021). University lecturers' beliefs and practices regarding task-based language teaching in Vietnam. *Education Sciences, 11*(11), 748. https://doi.org/10.3390/educsci11110748

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.

Ngan, D. T. N., Hercz, M., Khang, N. D., Thi Tuyet Van, P., Cang, N. T., Van Thong, N., & Le Bach Ngoc, P. (2025). Escaping the blame spiral: student-centered teaching reform in Vietnam. *Studies in Higher Education*, *50*(11), 2631-2645. https://doi.org/10.1080/03075079.2024.2444390

Ngo, C. L. (2024). Teacher agency for change and development in higher education: A scoping literature review. *International Journal of Educational Reform, 35*(3), 992-1015. https://doi.org/10.1177/10567879231224744

Nguyen, N. T., Baker, A., Eady, M., & Wright, J. (2023). Vietnamese EFL novice teachers' pedagogical decisions within a mandated communicative language curriculum. *The Journal of Asia TEFL, 20*(2), 338-356. https://www.doi.org/10.18823/asiatefl.2023.20.2.7.338

Nguyen, T. A., Le, T. T., Vang, M. D., Phuong, Y. H., Huynh, T. T. A., Nguyen, T. H., & Pham, T. T. (2023). Vietnamese EFL high school teachers' perceptions of difficulties when implementing competency-based English teaching curriculum and their proposed solutions. *Forum for Linguistic Studies*, *5*(2). https://doi.org/10.59400/fls.v5i2.1863

Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667336

Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury. https://doi.org/10.5040/9781474219426

Prime Minister. (2025). *Decision No. 2371/QĐ-TTg approving the Project on Transforming English into a Second Language in Schools for the Period 2025-2035, with a Vision to 2045*. Government of Vietnam. Retrieved from https://vanban.chinhphu.vn/?pageid=27160&docid=215728

Richards, J. C. (2006). *Communicative Language Teaching Today*. Cambridge University Press.

Tao, J., & Gao, X. (2021). *Language Teacher Agency*. Cambridge University Press. https://doi.org/10.1017/9781108916943

Vu, N. D. N., & Nguyen, N. Q. (2021). Task-Based Language Teaching to Enhance Learner Communicative Competence in Vietnam: A Matter of Opinion? *International Journal of Linguistics, Literature and Translation, 4*(8), 151-159. https://doi.org/10.32996/ijllt.2021.4.8.22

Willis, J. (1996). *A framework for task-based learning*. Longman.

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal, 37*(3), 308-328. https://doi.org/10.1177/0033688206071315

#115 CONCEPTUALIZATION OF TEACHERS' RESEARCH MINDSET AS A COMPLEX DYNAMIC SYSTEM IN THE ENGLISH LANGUAGE TEACHING CONTEXT

Hạnh Nguyễn Thị Minh[1], Xuan Mai Le[2], Trang Le Diem Bui[3]

Abstract

Research plays a crucial role in shaping the English Language Teaching (ELT) context, especially as demands for teacher professional development increase. This leads to a growing trend towards pursuing postgraduate programs that place greater demands on research competence. While studies on research are prevalent, a research gap remains regarding how ELT teachers – especially those who are graduates – conceptualize their research mindset. This study aims to fill that gap. This is Phase 1 of a larger study, utilizing a qualitative case study design incorporating a narrative frame (N=30) and in-depth interviews (N=4), which were conducted sequentially. Through the lens of Complex Dynamic Systems Theory (CDST), the results show that teachers' research mindset is not a static entity but a non-linear transition, from viewing research as a mandatory task to voluntary professional engagement. The study also highlights the paradoxes between practical perception, the self-efficacy gap, and diplomatic capability in the research process. The findings initially imply that integrating research into practice in the training program is necessary, as well as building scaffolding pathways to strengthen self-reliance and foster communities of practice to cultivate a sustainable research mindset among teachers.

***Keywords*:** research mindset, Complex Dynamic System Theory, ELT context, teacher professional development

1. Introduction

In the modern era of education, the role of English as a Language Teacher (ELT) is not simply that of a language transmitter; it has expanded to include that of a reflective practitioner and researcher within the classroom (Kamali et al., 2024). The teacher-researcher relationship has become one of the key elements for sustainable professional development. In line with this, a new relationship between teaching and research was proposed, and early studies (Brew, 2001; Hattie & Marsh, 1996) demonstrated the unity of research and teaching, grounded in the common pursuit of knowledge by teachers and students, with an increasing role. Hattie and Marsh's framework (1996) provided positive and negative relationships between teaching and research. This historical perspective supports the argument for a "teacher-researcher nexus," where roles are mutually reinforcing. Then, Brew (2001) posits that teachers are active decision-makers whose beliefs are deeply rooted in their own apprenticeship of observation as learners. This illustrates the early

[1] *Corresponding author, PhD student, School of Foreign Languages, Can Tho University, Vietnam; Email: hanhp2025001@gstudent.ctu.edu.vn*
[2] *Doctor, Head of the Department of General English and English for Specific Purposes, School of Foreign Languages, Can Tho University, Vietnam*
[3] *Doctor, Faculty of Foreign Languages at An Giang University, Vietnam National University, Ho Chi Minh City, Viet Nam*

acknowledgement of the teacher-research nexus, which lasts until current research concerns (Dao et al., 2022; Heng et al., 2022; Henry et al., 2020; Nguyen et al., 2016; Nguyen et al., 2022)

However, for a teacher to transition from a mere instructor to a researcher requires not only professional skills but also a certain orientation and intrinsic belief (Nguyen et al., 2022). While research in higher education is extremely important, many graduate students and faculty members still face significant psychological and institutional barriers. Although the concept of "mindset" as famously developed by Dweck (2006) also provides insight into why some people persevere in tackling difficult scientific problems while others give up, the specific concept of mindset about research in the ELT context is still scant. This study examines the role of developing a "research mindset" in graduate students using Complex Dynamic Systems Theory (CDST) to bridge the knowledge gap regarding intrinsic psychological components within the educational context of Vietnam.

2. Literature Review

2.1. Teacher research in English Language Teaching

Borg (2010) defines teacher research as a systematic, public inquiry aimed at the improvement of practice, drawing a clear boundary between engagement in research (production) and engagement with research (consumption). Research engagement is widely conceptualized as encompassing two interrelated but distinct forms: engagement with research (e.g., reading, consuming, and critically evaluating research) and engagement in research (e.g., designing studies, collecting and analyzing data, and disseminating findings) (Borg, 2010; Dao et al., 2022). Johnson và Hepworth (2022) also defined research in teaching as:

…research generated by practitioners working to understand their own practices and their own lives… begins with a puzzle or an issue related to practice that requires development and improvement (Johnson & Hepworth, 2022, p.31).

From the book of Banegas et al. (2022), the core characteristics of teacher research are performative and agentic. This positions teachers in an inquiry practice with an aim to bridge the gap between educational research and practice. Teachers are encouraged to utilize their knowledge and specific school surroundings to intervene in their teaching contexts (Banegas et al., 2022)

In this case, teacher research is viewed as a form of professional practice, not merely an academic product. In ELT, an inquiry-oriented mindset helps teachers view classroom problems as researchable issues, thereby fostering systematic reflection and evidence-based decision-making (Nguyen et al, 2022). Despite its value for professional development, direct research practice remains a minority activity due to institutional barriers and a technocratic orientation that views research as separate from teaching (Borg, 2010). The transition from consuming research to producing research is a developmental process, not an automatic one, and requires structured support such as mentoring and dialogue guidance (Dao et al., 2022; Nguyen et al., 2022).

2.2. A brief history of teachers' research mindset in English Language Learning

Mindset, mentioned as a target of concern to unpack individual differences, including capacity, cognition, and ethics (Dweck, 2006). In practice, individuals are capable of the capacity for lifelong learning and development through both nature (e.g., heredity or talent) and purposeful engagement (e.g., experience, training, individual effort, and environmental impact). From

Dweck's perspective, natural abilities serve as a starting point, and "purposeful engagement" facilitates the remainder of the process (Sternberg, 1985). Accordingly, learning and cognition occur interdependently. Elliot and Dweck (1988) developed "Achievement Goal Theory" to construct differences in response by distinguishing between performance goals and learning goals. Individuals seeking to validate their competence are considered to have performance-related goals; in contrast, those with learning goals maintain a perception of developing their current competence. To further understand the origins of these goals, the concept of mindset and theories of intelligence have been studied. Throughout extensive research on this issue, Dweck (2006) indicated that the perspective individuals apply to themselves profoundly affects their way of life, achievements, and goals. This opened a new research direction regarding the interdependence of beliefs, qualities, competence, and psychology, as well as their formation and whether they are fixed or malleable. From this, Dweck (2006) introduced the concept of mindset with two perspectives: fixed mindset and growth mindset. Specifically, those with a fixed mindset apply the concept of stasis to intelligence, while those with a growth mindset believe that intelligence is merely a starting point and is capable of changing over time.

Dweck (2008) identified two implicit theories of intelligence: "Incremental" and "Entity." First, aligned with the fixed mindset, entity theory posits that intelligence is a fixed and stable trait that humans possess from birth (Dweck et al, 1995). Individuals with this mindset tend to blame their own intelligence and ability when facing failure, which leads to the risk of poor performance. Conversely, those following the incremental theory possess a growth mindset, believing their intelligence can be molded, changed, and nurtured through effort and persistence; they are also ready to face novelty and challenges, desiring to develop competence and utilize their knowledge effectively (Dweck, 2000). In line with Achievement Goal Theory (Elliot & Dweck, 1988), mentioned above, growth-mindset individuals tend to set learning goals, whereas fixed-mindset individuals choose performance goals. When facing challenges, entity theory views them as threats, whereas incremental theory views them as information, including mistakes or failures (Dweck, 2008).

The analysis above demonstrates the connection between theories related to intelligence factors and learner achievement. It also shows the foundational nature of these theories for the terminology and classification of fixed and growth mindsets. Consequently, the term "mindset" is often cited as a foundational concept in theories of intelligence (Dweck & Molden, 2007). Overall, mindset is considered and researched as part of a belief system and self-regulation mechanism, providing a more comprehensive view of how mindset affects human motivation and behavior (Ryan & Mercer, 2012).

Mindset plays a crucial role in the learning process and learning outcomes. This is because learners' beliefs about their own abilities and intelligence significantly mediate how they approach learning and develop knowledge (Dweck, 2006). In other words, mindset can lead to differences in learning outcomes. Factors such as learners' beliefs, behaviors, needs, and skills are all related to their thinking systems (Dweck, 2006). Mindset in education is divided into two main types: fixed mindset and growth mindset. These two types of mindsets reflect an individual's belief that learning ability is fixed or can be developed through effort (Ryan & Mercer, 2012; Dweck, 2006). A fixed mindset causes learners to prioritize the validation of competence over its cultivation. Individuals with this mindset often avoid challenges, fear making mistakes, and are susceptible to disengagement when faced with difficulties (Dweck, 2008). A growth mindset encourages learners

to believe that abilities can be developed through effort and learning. This perspective is linked to incremental theory and learning goals, rather than focusing solely on results (Dweck, 2000; Elliot & Dweck, 1988). Thus, learners with a growth mindset are more likely to progress and develop academically than learners with a fixed mindset. This is because a growth mindset fosters motivation, perseverance, and a positive attitude towards mistakes. Learners view mistakes and feedback as opportunities to improve their learning (Dweck & Molden, 2007).

2.3. Teacher's research mindset as a continuous, dynamic, and integrated system under the complex dynamic perspectives

The construct of mindset has undergone a profound theoretical evolution over the last two decades. The analysis identifies Dweck (2006) as the predecessor of the field, establishing the influential dichotomy of fixed versus growth mindsets. This binary framework – categorizing individuals as possessing either fixed or growth mindsets – was instrumental in explaining motivational patterns, wherein individuals are categorized based on their belief in the malleability of human competence. However, a dominant theme emerging from recent scholarship is the problematization of this static framework.

The introduction of Complex Dynamic Systems Theory (CDST) as an explanatory mechanism is crucial here. The review highlights Larsen-Freeman (1997) as the initiator of this paradigm shift in language development, challenging linear models of mastery in favor of a process of becoming. Since language development is often dynamic and context-dependent, accompanying psychological mechanisms like mindset should also be viewed with corresponding flexibility and non-linearity. Therefore, current studies reframe mindset not as a static internal attribute, but as a dynamic psychological orientation that is emergent and context-dependent. From this point, mindsets are not merely possessed but are enacted through continuous interaction with social, institutional, and cultural environments. This change reflects a new trend in educational psychology, which focuses on the developmental process through experience and community collaboration, rather than viewing the research mindset as a fixed characteristic.

A research mindset can be examined as a dynamic and socially situated system of cognition and action toward research. Based on the conceptual framework with Complex Dynamic System Theory (CDST), the research mindset emerges not as a stable individual disposition but as a multilevel construct. Under the lens of CDST, this encompasses initial research experiences, cognitive mediation, behavioral enactment, and situational affordances.

2.4. Working definition of teachers' research mindset

The research mindset is not a static state or a rigid definition, but a developing concept with fluidity and high adaptability. Originating in science education (Wood, 2003) and adapted by Nguyen et al. (2022) for language teachers, this concept has continuously evolved through various stages (Lu et al, 2026). From focusing solely on the primary element of individual perception within the university environment (Wood, 2003), the concept has expanded to the field of language teaching to describe enduring individual dispositions (Taraban & Logue, 2012) closely linked to scientific understanding and readiness for action (Nguyen et al., 2022; Lu et al., 2026).

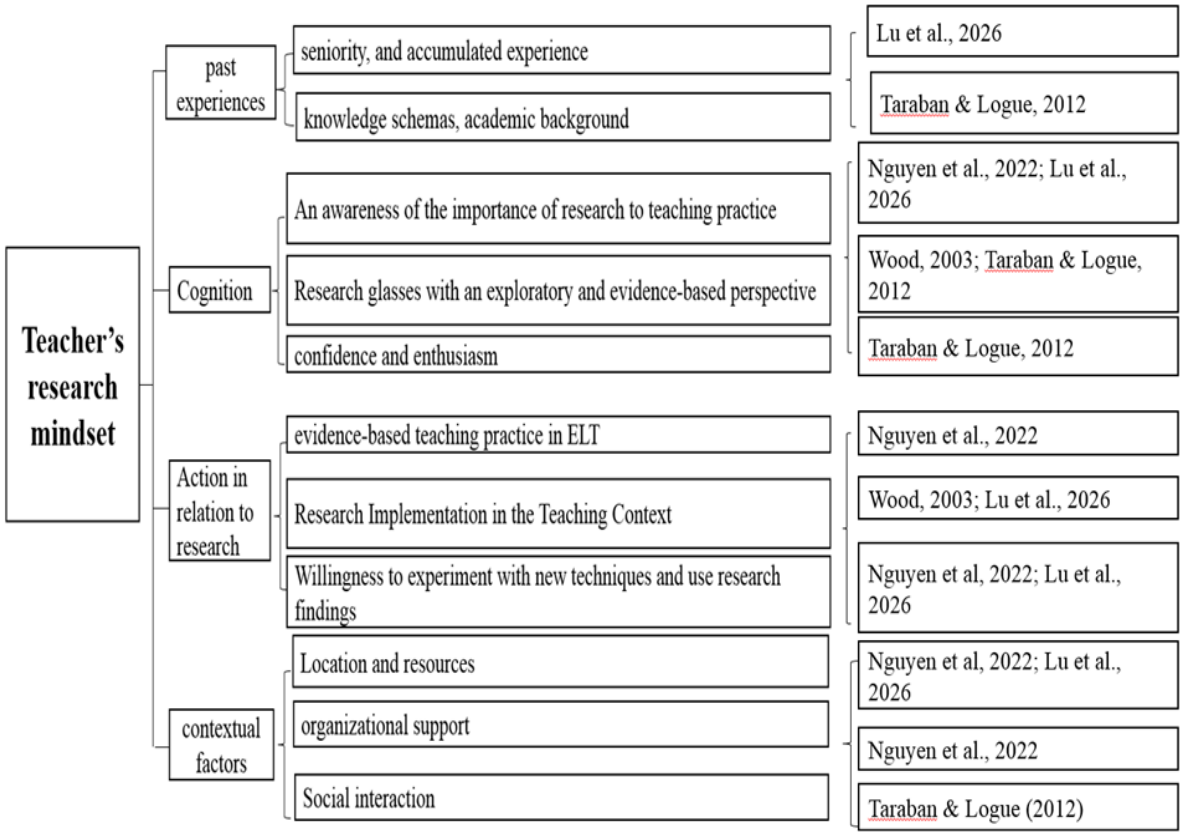
Because of its non-fixed nature, the research mindset continually opens new spaces for redefinition and enrichment across diverse professional contexts and empirical data. By placing the cognition and action elements of previous studies into the specific context of language teaching

education, Lu et al. (2026) emphasized the importance of recognizing the teacher's dual roles as both practitioners and researchers. This aims to strengthen the link between research and practice. This perspective aligns with the current needs for teacher development in language teacher education (LTE) "for teachers to engage not only with research to inform their teaching …but also to carry out research in their classrooms" (Lu et al., 2026). Based on this, the present study applies the perspective of language teachers as both practitioners and researchers to explore the characteristics of their thinking and research practices when assuming this dual role, and how they develop and strengthen the relationship between research and practice.

Based on the development process of the research mindset concepts mentioned above, in this study, teachers' research mindset in ELT is operationalized as a persistent yet adaptive disposition toward evidence-based problem-solving, encompassing their perceptions and actions related to research in ELT. It functions as a dynamic system wherein teachers' cognition (an awareness of the importance of research to teaching practice, research glasses with an exploratory and evidence-based perspective, confidence, and enthusiasm) and actions (evidence-based teaching practice in ELT, research implementation in the teaching context, willingness to experiment with new techniques and use research findings) related to research in ELT are mutually constitutive and continuously mediated by past experiences (knowledge schemas, seniority, accumulated experience, and academic background) and situated professional context (location and resources, organizational support, social interaction).

Figure 1

Components of the teacher's research mindset



2.5. Current studies on the teacher research and teachers' research mindset

Although the relationship between research and teaching is widely acknowledged, much of previous research has focused on the influence of external factors on research productivity or the role of research in professional development (Hattie & Marsh, 1996). In recent years, a significant shift has been observed as research has pivoted toward internal factors and the process of fostering a research mindset (Smith & Rust, 2011; Wilson et al., 2017; Dao et al., 2022; Nguyen et al., 2022; Truong et al., 2025).

These studies have shown that pedagogical interventions at the micro level can promote research thinking, particularly through research-based learning, fostering critical thinking, academic engagement, and a learning community among students and faculty (Smith & Rust, 2011). First, research programs for students, such as the Summer Undergraduate Research Fellowship, demonstrate that direct participation in research can enhance problem-solving skills, academic understanding, and learning motivation (Wilson et al., 2017). Additionally, teacher-learner engagement in research requires structured pedagogical support, such as the Research Engagement Process (REP) model, which utilizes scaffolding and expert guidance to help learners transition from reading research to conducting research (Dao et al., 2022). Practical research experiences and community practice also play a crucial role in developing inquiry-oriented thinking and research identity through academic collaboration and shared problem-solving (Nguyen et al., 2022; Truong et al., 2025).

However, other studies have also shown that the development of a research mindset at the individual level does not necessarily translate into research productivity at the organizational level, as it is influenced by structural constraints within the higher education system (Harman & Ngoc, 2010; Fatseas, 2010; Heng et al., 2022; Nguyen et al., 2016). More specifically, institutional factors and working conditions such as academic qualifications, income, and teaching workload significantly influence faculty research productivity (Henry et al., 2025). In the Vietnamese context, although the government has promoted higher education reform and strengthened research capacity, universities still face many limitations, such as a lack of resources, research infrastructure, and research culture (Harman & Ngoc, 2010; Fatseas, 2010). High teaching workloads and financial pressures lead many faculty members to prioritize teaching over research, resulting in low scientific publication rates at regional universities (Nguyen et al., 2016). Similar challenges are also observed in the Southeast Asian context, where a gap exists between policy discourse and academic practice, even though research and teaching are often seen as complementary (Heng et al., 2022).

2.5. Problem statement

Overall, evidence suggests that the development of research-oriented thinking is a result of the interaction between micro-interventions and macro-structural conditions, highlighting the need to consider both levels simultaneously when studying research engagement in higher education. This reveals research gaps in focusing solely on surface productivity, the dissonance within the research-teaching nexus, and the need for a deeper understanding of how the research mindset of teachers has been defined, and how this understanding should be changed or maintained within the education system and postgraduate training environment.

2.6. Research aim and research question

Acknowledging the gap mentioned above, the researcher conducted an exploratory sequential longitudinal study, which aims to not only explore teachers' conceptualization of their research mindset but also examine whether their research mindset changes after completing an official post- graduate program, their Master program. This study consists of two stages with four different phases, conducted sequentially.

In this paper, we present the first phase of Stage 1 of this study, which aims to explore the conceptualization of teachers' research mindset with the specific question:

How do teachers, as graduates, conceptualize their research mindset during their Master's program?

3. Methodology

3.1. The design

In this current study, we employed a qualitative case study design, as the first phase in stage 1 of another exploratory sequential longitudinal mixed-methods study.

The narrative frame was used in this to serve a large cohort, allowing broad access to participants' reflections while maintaining a structured format. Then, in-depth interviews were conducted to gather insights about how they conceptualize their research mindset in the ELT context.

Narrative frames are recognized as a legitimate and increasingly established method in narrative inquiry within EFL and teacher education research (Wicking, 2021). This was first developed by Barkhuizen and Wette (2008) and is defined as a written story template consisting of a series of incomplete sentences and blank spaces of varying lengths (Barkhuizen, 2011). The collection of rich qualitative data from a large number of participants without the constraints of time, geographical distance, or scheduling that typically limit face-to-face data collection (Barkhuizen, 2014). Methodologically, the structured nature of narrative frames helps participants organize their reflections logically and coherently, which is especially valuable when working with respondents with varying levels of academic writing proficiency (Wicking, 2021). Moreover, completing the frames independently, without direct researcher presence, may reduce social desirability bias and encourage more candid responses. Nevertheless, narrative frames also present certain limitations. Participants may perceive the structured format as restrictive or frustrating, and the absence of direct interaction prevents researchers from clarification or elaboration by asking any follow-up questions (Wicking, 2021). Narrative frames were not used as the sole data source (Barkhuizen, 2014) but were complemented by in-depth interviews. In this study, narrative frames therefore functioned both as an initial exploratory tool and as a foundation for the subsequent in-depth interviews (Creswell & Guetterman, 2024). Thus, this is considered appropriate in the study to strengthen the protocols for the following in-depth interview and to overcome the limitations of a small participant sample.

As mentioned in the section on the working definition of teachers' research mindset in ELT, some empirical studies contributed to the development of teachers' research mindset (Woods, 2003; Taraban & Logue, 2012; Nguyen et al., 2022; Lu et al., 2026). Narrative frames and in-depth interview protocols were designed based on findings of these empirical studies on

teachers' research mindset with four main clusters. The four clusters consist of teachers' previous experiences, their cognition, their action in relation to research, and the contextual factors (See Figure 2.1).

3.2. The context and participants

3.1.1. The context and participants

Participants are teachers in English Language Teaching and are studying in a Master's program at a university in the Mekong Delta. They were recruited for many reasons. Firstly, due to the need for teacher professional development up to 2030 (MOET, 2021), an increasing number of teachers, especially core teachers at schools, institutions, or universities, are participating in Master's programs. Given their current teaching and research roles, as well as their role as future human resources, their research mindset at this stage is extremely important. Secondly, in terms of the context, choosing a Master's program at a university in the Mekong Delta is for the following reasons. This is not just about enhancing knowledge, but the most crucial transition period for a teacher to become a researcher. Furthermore, this is one of the region's prestigious and important institutions for training postgraduate human resources. Finally, in the Vietnamese context, the diversity of input (those who have previously completed a graduate thesis and those taking a substitute course) creates a significant gap in readiness. This research is crucial because it helps determine how this difference in background affects the ability to absorb the research mindset during the two years of training.

3.1.2. Sampling

More than 30 participants were recruited by using convenience sampling. They were invited to complete a narrative frame in person. This sampling takes advantage of the availability and willingness of Master's students who are studying in the same university.

The participants have finished the three courses in relation to educational research, including courses for research methodology, data collection and analysis, and writing for publication. At the end of these courses, each of them had a proposal and some preliminary data as course outcomes and preparation for their graduation thesis. All of the invited participants passed these courses with an acceptable quality of outcomes.

Based on the results of the narrative frame, which showed their in-depth response and willingness to participate in the next phases, four of the 30 participants (detailed information is presented in Table 1 below) were selected to individually join an online in-depth interview, using purposive sampling. We also tried to ensure that the four participants who were invited also had diverse academic and teaching backgrounds, from a novice freelance tutor to a senior teacher in a public high school.

Table 1

Participant description

No.	Coded Name	Description
P1	Hoa	She is an English teacher at an English language center.

		Previously, she mainly learned about research through documents, books, and articles. She had never participated in any research.
P2	Huy	He is a senior English teacher at a High school. He has never conducted research, only knowing about it through two courses in his Bachelor's program.
P3	Nam	He is a teacher at an English tutoring center. He habitually reads short articles about research in various fields from sources such as National Geographic. In addition, he has participated as a participant in research. From this, he has an understanding of research but has never conducted any research.
P4	Tâm	He is currently not officially employed by any organization. He is pursuing a Master's program immediately after graduating with a Bachelor's degree. Previously, he had experience approaching research from courses at the Bachelor's level and tried applying it to research during his Master's program.

3.2. Data collection and analysis procedure

Data collection

After being designed, the template of narrative frames was piloted with at least two students (different from the participants of this stage) to test its effectiveness and receive constructive feedback. The results were collected and compared with the study objectives, and the researcher and supervisor carefully considered constructive feedback to rewrite and simplify some sections before sending them to participants. After completing narrative frames, 4 participants were selected and invited to an in-depth interview. Each interview was conducted individually in approximately 90 minutes.

Data analysis

Thematic analysis is a method developed by Braun and Clarke (2006), primarily suitable for qualitative data in psychology, social science, or education. This method can not only organize and describe the data set in detail but also interpret various aspects of the research topic (Braun & Clarke, 2006). Thematic analysis (Braun & Clarke, 2006) fits this design because of its flexibility to identify broad patterns in how participants conceptualized teachers' research mindset during the program.

The analysis process of the data includes six steps. After gathering narrative frames from target participants, I familiarized myself with the data by lining up and rearranging. Then, I generated initial codes after grouping, reading, and analyzing raw data. The third step helped search for themes, and then reviewed them again and again in the next step. Themes, latter, were refined and named. Finally, the report was produced with generated themes and sufficient excerpts. Noticeably, I also asked for my lecturers' and supervisor's checks for both the generated codes and themes. These themes played a crucial role in the revision of the in-depth interview protocol, in-depth interview participant recruitment (mentioned in the Sampling section), and data triangulation. Then, after gathering in-depth interview data, this six-step process was conducted again, also under the supervision of the lecturers and supervisors, to analyze the in-depth interview data. The two sources of data were then triangulated to find out the final themes.

4. Results

The research results show that, although all four participants were enrolled in the same Master's degree program, differences in background, research experience, and current work environment led to differences in how they conceptualized their research mindset. This demonstrates that a research mindset is not a fixed concept but a dynamic system that is constantly changing and influenced by both experience and the current environment.

Theme 1: The research mindset is about continuous improvement to solve real-world problems

Hoa conceptualized her research mindset as being closely linked to innovation and improving the quality of teaching through reflection, inquiry, and independent research. She did not view research as a "distant theoretical field" but rather as a continuous mindset process aimed at finding solutions to problems, especially those in her own classroom. Hoa said:

"I think a research mindset is about continuous improvement and finding new results…it's a system in which I constantly review my teaching process…to solve problems and improve quality."

Regarding her perspective on research and teaching, Hoa previously felt that research did "not support teaching much because it was time-consuming". However, she realized there is a "strong connection when research results can support practical work," and conversely, a "separation" if it is "merely paper theory." Although Hoa currently identifies herself primarily as someone who "uses research for her graduate studies", she believes in her ability to develop her research competence in the future and affirms that research is necessary.

Linking with Hoa's teaching context and previous experiences, reported in Table 1, and her narrative frame, her research mindset is conceptualized as a tool for classroom improvement, which needs to be practical. This stemmed directly from her purely theoretical experiences from books, which caused the initial confidence gap in her research mindset. Her teaching context at a language center, which is mostly product-driven and demands tangible results, also affected her research mindset and characterized it with innovation and improvement in her teaching quality, with direct and practical effects in her classrooms.

Sharing the same conceptualization approach, Huy has taken the research mindset beyond the realm of knowledge or skills, instead focusing on attitude and emotion. He conceptualized the research mindset as a continuous process of learning and self-evaluation. He said:

"A research mindset includes learning, reflection, application, and self-evaluation."

But he added the element of "curiosity and dedication to discovering knowledge." He believes that a teacher needs "an open-minded, inquisitive, and inquisitive mindset." Like Hoa, for Huy, research must have "practical value and help teachers develop their problem-solving thinking." However, he faces "time pressure," especially when balancing teaching and participating in a Master program, which sometimes weakens his thinking. Nevertheless, Huy remains "determined to continue" his research in the future to bring "value to learners and the community."

Mentioning curiosity and dedication as characteristics of Huy's research mindset reflected the interaction between his seniority and environmental pressure. As a high school teacher, he also

lacked practical research experience like Hoa and experienced the workload at the public school, which limited his research engagement. This acts as a strong external constraint and leaves his research mindset heavily influenced by his own determination, reflected in his internal goal to bring value to his community.

Theme 2: Research Mindset is the ability to think systematically and logically, inspired by practical values

Nam believed that a research mindset is a process of logical thinking, where teachers "do not perceive, evaluate, and solve problems emotionally or formally." He emphasized structure and observation skills to adjust teaching methods. He defined it as:

"A research mindset is the ability to think systematically, logically, and flexibly in problem-solving… helping me approach research problems systematically and avoid formalistic research."

However, he often encountered a conflict when conducting research "does not fit within his time constraints and heavy teaching workload at school." He felt "lacking confidence in his ability to conduct systematic research" and "does not believe" he "can further develop this ability" due to "lack of opportunities for training." Because he prioritized teaching, Nam plans "not to continue doing research in the future."

Nam's research mindset showed a focus on a logical and systematic approach, shaped by his habit of reading multidisciplinary scientific papers and his experience as a research participant. This experience helped him develop a knowledge schema. Even though he did not directly conduct his own study, as a consumer and participant of educational research, he exposed and understood the nature of research. However, his current teaching context with teaching intensity and a lack of training opportunities, as he mentioned above, hindered his faith in his potential to conduct research independently.

Tam also shared this view, but took a more practical approach. He conceptualized that "a research mindset must be based on proven evidence and is part of professional development." He stated:

"A mindset is something sustainable over time… based on existing research, because it has been validated. A teacher with a research mindset is someone who can utilize existing research as well as obtain data themselves within the current teaching context."

Tam used to think "research is unnecessary for teachers," but now he believes it is part of "being a professional teacher." He particularly saw a "strong connection" in the action research model. He also emphasized the "support of modern tools like AI in developing research capabilities". Regarding research engagement, he "affirms the possibility" of participating in research in the future. However, he was "not entirely sure" about whether to "conduct the research [himself] or simply use" existing research. He also stated that he "would not commit to a long-term engagement unless it was absolutely necessary for his studies or work."

Tam demonstrated a pragmatic and professional research mindset, reinforced by his current academic context as a full-time Master students. Despite following the same Master program as the three former participants, he showed his background acted as an ideal catalyst with no workload pressure. Lacking current teaching experience, his research mindset was stuck with an academic

foundation in his Bachelor program and his self-study using AI and other supporting tools. However, due to his limited teaching practice, he cannot make sense of the role between teaching and research, and showed his uncertainty with his future commitment to do research.

5. Discussion

5.1. Unpacking their conceptualization of research mindset with Complex Dynamic System Theory

Under the lens of Complex Dynamic System Theory (CDST) (Larsen-Freeman, 1997; Larsen-Freeman & Hiver, 2025), a research mindset is viewed not as a static trait, but as an emergent, non-linear, and socially situated system. Using the above-identified themes, this study can briefly unpack this system through its interacting components, constraints, and attractors. Furthermore, based on the themes above, a person with an evolving research mindset in an ELT context can be described in the following paragraph.

Environmental factors act as "attractions/constraints" shaping systemic behavior. The large workload and time scarcity are strong constraints limiting research participation. Research is difficult to become a regular activity because the time and effort costs (thinking, completing research) are too high. The difficulties of workload, time scarcity, and the lack of connection between research and classroom reality shared by students have been confirmed in recent studies by Henry et al. (2020) and Heng et al. (2022). CDST views these as strong constraints hindering the development of mindset systems. The study by Lu et al. (2026) adds that teaching location (urban and rural) also creates significant differences in cognition and action due to disparities in resources and cultural support for research. This means that the research mindset may develop gradually with effort and time, but competence needs time and conditions. Therefore, without environmental intervention, such as support from the school or training program, persistently low research engagement or experiences can lead to a decline in research mindset over time, as mentioned in the study by Taraban and Logue (2012).

Then, non-linearity and transformation in the training process also occur. According to CDST, the system changes when there is a catalyst. Master's programs act as a catalyst, helping learners shift from being consumers of research to having a more multi-dimensional understanding of research (Dao et al, 2022). Further, a study by Lu et al. (2026) points to a paradox that novice teachers often have a very high cognitive mindset but a lower action-oriented mindset compared to experienced teachers. This reflects the non-linear shift observed in current findings when learners begin to move from viewing research as a mandatory task to a voluntary professional commitment when they realize the practical value in solving classroom problems. This shift demonstrates the emergent nature of the teacher-researcher identity thanks to the training process. Therefore, situational factors are both hindrances (teaching pressure, fixed curriculum) and motivators (passion for discovery, Master's program).

In terms of attractor state and professional identity, the majority of participants position themselves as users or consumers of research rather than researchers. The main reason is a lack of confidence in their research abilities (e.g., confident in reading literature but lacking the ability to identify a topic or conduct methodological research) (Nguyen et al., 2022). This causes the thinking system to be "stuck" in the consumer state. Behavior and feedback loops are considered within the system. When faced with difficulties, learners seek social support (colleagues, experts).

They receive feedback – reflect – adjust their goals and methods, demonstrating self-regulation and adaptability. The prominent professional identity is that teachers are always thinking, curious, and seeking new solutions to improve teaching (Wood, 2003). This is driven by critical thinking and long-term goals aimed at providing value to learners.

The cognitive component of the research mindset is also presented. Participants see themselves as teachers with critical thinking skills, but still primarily as consumers of research. They value the practicality of research, viewing it as a tool for solving classroom problems. A lack of confidence in their abilities prevents them from taking on the role of a researcher. Then, about the behavioral component, believing that research skills can be developed through learning and practice. Demonstrating perseverance, seeking literature and support from professional networks to improve their abilities.

In conclusion, the research mindset of teachers participating in this study is pragmatic, highly dependent on social context, and currently in a state of academic transition. This state is facilitated with a growth orientation, when the perception of research shifts from a forced requirement to a necessary professional investment with personal traits like curiosity and passion.

5.2. Research-teaching paradox and transition emerging from teachers' research mindset

The above results show similarities with Borg's (2010) study, where viewing oneself as merely a consumer of research products was mentioned as one of the two main types of teacher engagement with scientific research in the field of education. Furthermore, Nguyen et al.'s (2022) study found that while teachers had graduated with the necessary knowledge and understanding for research, they lacked confidence in their ability to conduct research and had not had much practical experience. The study of Priyadarsini and Kumar (2024) emphasizes that practical research experiences (UREs) and effective mentoring are key catalysts for improving self-efficacy and fostering a research mindset. This explains why Hoa and Huy lack confidence in their own research capabilities, regardless of their trained research skills. On the contrary, Clark and Johnstone's (2018) study of undergraduate students also indicates that overconfidence sometimes masks actual shortcomings in scientific research and citation skills. This seems different when considering teachers as graduates, as they tend to underestimate their scientific research and citation skills and overestimate the difficulty of the research procedure.

The lack of a tendency to participate in research due to a preference for teaching demonstrates a negative correlation, as seen in Hattie & Marsh's study (1996). Simultaneously, the difficulties they shared regarding workload, research not being closely aligned with reality (Ding et al., 2025), or lack of time have also been confirmed in recent studies (Henry et al., 2020; Heng et al., 2022). The current research emphasizes the deep link between research and problem-solving ability, a perspective also found in high school students in the study by Ding et al. (2025), where research is understood as a continuous process of curiosity, progression, and in-depth exploration. For teachers like Nam, who prioritize teaching over research is not only a personal choice but also a consequence of institutional factors such as assessment standards focused on course overall scores rather than research outcomes. Their denial of either the research or teaching role, despite recognizing the relationship between the two, shows the variability of the research mindset when environmental factors become unfavorable. These factors can include institutional and working conditions such as educational qualifications, income, and teaching workload (Henry

et al., 2020), or a lack of resources, research infrastructure, and research culture (Harman & Ngoc, 2010; Fatseas, 2010).

Additionally, the belief that one cannot develop research capabilities or that one's mindset is always fixed also reflects a fixed mindset (Dweck, 2006). In this case, all past and present difficulties and failures become barriers to future research participation. To overcome the problem of a fixed mindset, the role of Communities of Practice is essential (Dao et al., 2022). Walker et al. (2020) pointed out that collaboration among colleagues not only creates a safe space for experimentation but also helps teachers overcome feelings of isolation and build researcher identity. Conversely, there is still a belief that one's own ability can be developed, which reflects a growth mindset (Dweck, 2006). In the case of Hoa and Huy, their belief in the potential for future research capacity development is an expression of a growth mindset according to Dweck (2006), which helps them view current challenges as opportunities and affirm their intention to continue participating in research in the future. The fact that students in this study proactively sought support from the professional network when facing difficulties is a positive feedback loop, helping the mindset system to self-adjust and develop sustainably (Wilson et al, 2017; Nguyen et al., 2022). The overview of teachers' research mindset emphasizes the connection between research ability and inquiry, a factor nurtured through training as participants gradually apply their studies to teaching. They see a deep link between research and their problem-solving abilities, as demonstrated in previous studies (Wilson et al., 2017) and with teaching (Nguyen et al., 2022; Truong et al., 2025).

6. Conclusion

This study has provided a comprehensive picture of the research mindset of graduate students majoring in English Language Teaching (ELT), affirming that it is not a static psychological trait but a dynamic, non-linear system that emerges from the complex interaction between the individual and the environment. The research results point to three key points in the formation of this mindset.

The results indicate a clear shift in teacher motivation and identity while revealing the persistent paradoxes between perception and actual competence. The former shift is characterized by a transition from viewing research as a mandatory academic obligation at the undergraduate level to a voluntary professional commitment at the graduate level. The Master's program acts as a crucial catalyst, helping students move from the position of passive knowledge consumers to multi-dimensional research users, beginning to form the identity of a teacher-researcher. In contrast, the latter paradox persists between perception and intrinsic ability. There exists a gap between appreciating the practical value of research and believing in one's own research capabilities. Although students possess a positive research mindset and a desire to improve teaching, they remain stuck in the attractor state of a consumer due to a lack of confidence in methodology and pressure from external barriers such as limited time and heavy workload. This leads to the third point – adaptability and nurturing a research mindset. Under the lens of CDST, teachers' research mindsets are driven by flexibility through social adjustment strategies (like seeking support from colleagues and the use of digital tools). Their perseverance and emotional resilience in the face of difficulties show a growth mindset, suggesting potential for ongoing system evolution.

These findings showed some implications. From the perspective of CDST, the research mindset of teachers is constantly evolving and influenced by their surrounding context. Therefore, after training and acquiring theoretical knowledge of research, they showed a shift. Cognitively, they viewed research as an ongoing process of professional engagement rather than an imposition. In terms of action, it transitions individuals from inactive engagement to autonomous inquiry. This shows that, due to its non-linear nature, the research mindset can be influenced by its surroundings, specifically the training environment. Furthermore, teachers only shift from passive to active engagement and take practical action when they perceive a connection between research and teaching. Therefore, to sustain this non-linear development, designing a training program aligning with individual research topics and teaching practices is crucial. When teachers actively engage in research, their cognition and actions interact, helping them deepen their theoretical understanding and improve their practical skills. In this case, the bidirectional interaction between cognition and action will be directly enacted.

Furthermore, although they all undergo the same training program, not all teachers believe they will be teacher-researchers in the future. This highlighted the importance of external factors, including scaffolding methods, supervisors, and mentoring. A lack of self-confidence, even after completing part of the training process, can underscore an insufficiency of isolated structural training. Because the research mindset is a continuous and dynamic process, it demands more than a formation stage, but a systematic ecological nurturing through robust communities.

In short, the research mindset of ELT teachers as graduates is currently situated in an academic transition. To foster the sustainable development of their research mindset, training programs need to focus more on bridging the gap between theory and practice, while establishing a supportive institutional context to help teachers overcome self-reliance barriers, thereby nurturing and maintaining their research mindset for further potential of research engagement within their professional context.

However, there is a limitation on participants who can only represent the current Master training program at a university, while a research mindset appears in many other contexts with different subjects, like novice-teachers, university students, etc. Therefore, further studies can widen the scope or look at different subjects. Additionally, the current results can only present the first phase of a longitudinal study, so they can merely capture teachers' research mindset after their in-class courses with limited real research experiences. This leaves room for further studies to clarify whether their research mindset changes after they finish their graduation thesis and return to the teaching context.

References

Banegas, D. L., Edwards, E., & de Castro, L. S. V. (Eds.). (2022). *Professional development through teacher research: Stories from language teacher educators* (Vol. 102). Multilingual Matters.

Barkhuizen, G. (2011). *Home tutor cognitions and the nature of tutor-learner relationships. In P. Benson, & H. Reinders (Eds*.). Beyond the language classroom (pp. 161-174). London: Palgrave Macmillan.

Barkhuizen, G. (2014). Revisiting narrative frames: An instrument for investigating language teaching and learning. *System, 47*, 12-27

Barkhuizen, G., & Wette, R. (2008). Narrative frames for investigating the experiences of language teachers. *System, 36*(3), 372-387.

Borg, S. (2010). Language teacher research engagement. *Language Teaching, 43*(4), 391–429. https://doi.org/10.1017/S0261444810000170

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brew, A. (2001). Conceptions of research: A phenomenographic study. Studies in Higher Education, 26(3), 271–285. https://doi.org/10.1080/03075070120076255

Clark, J.C., & Johnstone, J.L. (2018). Exploring the Research Mindset and Information-Seeking Behaviors of Undergraduate Music Students. *Coll. Res. Libr*., 79, 499-516. https://doi.org/10.5860/crl.79.4.499

Creswell, J. W., & Guetterman, T. C. (2024). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson

Dao, P., Nguyen, M. X. N. C., Iwashita, N., & Spinelli, F. (2022). A dialogic approach to fostering TESOL teacher-learners' research engagement: Insights from a "learning-to-research" procedure. *TESOL Quarterly, 56*(2), 775–787. https://doi.org/10.1002/tesq.3127

Ding, L., Sherafat, A., Zhang, J., Underwood, T., Guo, D., Muirhead, J., ... & Kerr, J. (2025). Cultivate a Research Mindset: How Do High School Students Understand and Approach Authentic Research?. In *Proceedings of the 18th International Conference on Computer-Supported Collaborative Learning-CSCL 2025, pp. 178-186*. International Society of the Learning Sciences.

Dweck, C.S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: a world from two perspectives. *Psychological Inquiry*, *6*, 267-285.

Dweck, C.S. (2000). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Taylor and Francis. Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Dweck, C.S. & Molden, D. (2007). Self-theories: Their impact on competence, motivation and acquisition. In A. Eliot and C. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 122-140). London: The Guilford Press.

Dweck, C. S. (2008). *Mindsets and math/science achievement.* New York: Carnegie Corporation of New York.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of personality and social psychology*, *54*(1), 5.

Fatseas, M. (2009). Research–industry cooperation supporting development in Vietnam: The challenge of translating policy into practice. In *Reforming higher education in Vietnam: Challenges and priorities* (pp. 103-115). Dordrecht: Springer Netherlands.

Harman, G., & Ngoc, L. T. B. (2009). The research role of Vietnam's universities. In *Reforming higher education in Vietnam: Challenges and priorities* (pp. 87-102). Dordrecht: Springer Netherlands.

Hattie, J., & Marsh, H. W. (1996). The Relationship Between Research and Teaching: A Meta-Analysis. Review of Educational Research, 66(4), 507-542. https://doi.org/10.3102/00346543066004507

Heng, K., Hamid, M. O., & Khan, A. (2022). Academics' conceptions of research and the research-teaching nexus: Insights from Cambodia. *International Journal of Educational Development*, *90*, 102569.

Henry, C., Ghani, N. A. M., Hamid, U. M. A., & Bakar, A. N. (2020). Factors contributing towards research productivity in higher education. *International Journal of Evaluation and Research in Education, 9*(1), 203–211. https://doi.org/10.11591/ijere.v9i1.20420

Johnson, N., & Hepworth, M. (2022). Ghosts in the machine? Exploratory teaching on a distance learning development project. In D. L. Banegas, E. Edwards, & L. S. Villacañas de Castro (Eds.), *Professional development through teacher research: Stories from language teacher educators* (pp. 31–50). Multilingual Matters. https://doi.org/10.21832/BANEGA7710

Kamali, J., Freeman, D., Larsen-Freeman, D., Norton, B., & Farrell, T. S. C. (2024). Editorial: Language teacher education research – key trends, challenges, and questions. *Language Teacher Education Research, 1,* 1–16.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 141–165.

Larsen-Freeman, D., & Hiver, P. (2025). Complex dynamic systems theory. In *Theories in Second Language Acquisition* (3rd ed.). Routledge.

Lu, Q., Yao, Y., Xiao, L., & Zhu, X. (2026). Research mindset and practices in classroom assessment among secondary school EFL teachers: Characteristics, relationships, and strategies to strengthen their connection. *Teaching and teacher education*, *170*, 105282.

Ministry of Education and Training. (2021). Circular No. 25/2021/TT-BGDĐT: Guiding the implementation of doctoral and master's level training for lecturers at higher education institutions under Decision No. 89/QD-TTg.

Nguyen, M. X. N. C., Dao, P., & Iwashita, N. (2022). Nurturing teachers' research mindset in an inquiry-based language teacher education course. *The Modern Language Journal, 106*(3), 599–616.

Nguyen, Q., Klopper, C., & Smith, C. (2016). Affordances, barriers, and motivations: engagement in research activity by academics at the research-oriented university in Vietnam. *Open Review of Educational Research, 3*(1), 68–84.

Kumar, S. P. (2025). Nurturing research mindset through undergraduate research experience (URE) and mentoring in engineering institutions. *Journal of Engineering Education Transformations*, 92-100.

Ryan, S., & Mercer, S. (2012). Implicit theories: Language learning mindsets. In *Psychology for language learning: Insights from research, theory and practice* (pp. 74-89). London: Palgrave Macmillan UK.

Smith, P., & Rust, C. (2011). The potential of research-based learning for the creation of truly inclusive academic communities of practice. Innovations in education and teaching international, 48(2), 115-125. https://doi.org/10.1080/14703297.2011.564005

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of personality and social psychology*, *49*(3), 607.

Taraban, R., & Logue, E. (2012). Academic factors that affect undergraduate research experiences. *Journal of Educational Psychology, 104*(2), 499–514.

Truong, T. P., Le Bao, N. T., Nhu, Y. V. N., Kieu, D. N. N., & Thanh, D. T. (2025). Exploring English Studies Majors' Interactions in Collaborative Research Teams as Communities of Practice. *Educational Process: International Journal*, *18*, e2025424.

Walker, L., Carson, M., Datti, D., & Lee, J. (2020). Joining the club: Teachers collaborate to develop a research mindset and unlock the benefits of teacher research. *Teacher Development Academic Journal, 1*(1), 66-83

Wilson, J., Wu, Y., Xie, J., Johnson, D., & Huijser, H. (2017). Inspiring learning through research and enquiry: The Summer Undergraduate Research Fellowship (SURF) at Xi'an Jiaotong-Liverpool University (XJTLU). *Developing the higher education curriculum: Research-based education in practice*, 46-59.

Wicking, P. (2021). *Qualitative content analysis for describing and interpreting narrative frames.* In P. Wicking (Ed.), Applied Linguistics and Foreign Language Education Research - A Vision for the Future (pp. 230-244). Kinseido

Wood, W. B. (2003). Inquiry-based undergraduate teaching in the life sciences at large research universities: A perspective on the Boyer Commission Report. *Cell Biology Education, 2*(2), 112–116.https://doi.org/10.1187/cbe.03-02-0004

Bionote

Thi Minh Hanh Nguyen, Can Tho University

Nguyen Thi Minh Hanh is a PhD student in Principles and Methods in English Language Education at Can Tho University, Vietnam, where she was honored to be admitted directly to the doctoral program without completing a Master's degree. In recognition of her early research contributions, her previous work has focused primarily on understanding group work dynamics in language education. Building upon these experiences, Hanh's current doctoral research explores teachers' research mindsets in English language teaching practices, with the hope of contributing to professional development practices in Vietnamese higher education.

Mai X. Le, Can Tho University

Dr. Le Xuan Mai, Head of the Department of General English and English for Specific Purposes, School of Foreign Languages, Can Tho University, Vietnam, obtained her doctorate degree in Education in Australia. Her research interests include English language teaching (ELT), ICT in ELT, teacher education, and teacher professional development.

Trang Le Diem Bui, An Giang University, Vietnam National University Ho Chi Minh City

Dr. Bui Le Diem Trang is a senior lecturer in the Faculty of Foreign Languages at An Giang University, Vietnam National University, Ho Chi Minh City. She has a master's degree in Applied Linguistics from the University of Queensland, Australia, and a PhD in Applied Linguistics from Victoria University of Wellington, New Zealand. She has over 15 years of teaching experience, specializing in training pre-service and in-service EFL teachers. Her professional interests include English teacher education, task-based language teaching (TBLT), and technology in language education.

#116 THE EFFICACY AND STUDENT PERCEPTIONS OF AI-GENERATED VERSUS TEACHER FEEDBACK ON UNIVERSITY STUDENTS' L2 WRITING DEVELOPMENT: A MIXED-METHODS STUDY

Anh Tran Duc[1]

Abstract

The arrival of GenAI tools in language education has necessitated an understanding of their efficacy, particularly in propelling learners to greater writing proficiency. This study seeks to explore the extent to which Automated Writing Feedback (AWF) generated by ChatGPT can assist learners in writing, together with their perceptions of its effectiveness, compared to the traditional teacher-led method. This 10-week, mixed-methods research investigated two groups of 54 students who either received feedback from the educator or from AI during completion of 8 paragraph writing tasks. Findings retrieved from quantitative data (pre-test/post-test scores) indicate improvements in writing scores of both groups, while qualitative data (survey results and interview responses) suggest that teacher feedback played a meaningful role in fostering students' affective aspects. We conclude that AI holds the potential to assist the marking process by handling micro-level errors, supplying timely feedback, allowing educators to prioritize macro-level errors, and providing a more thorough feedback experience. Thus, this paper proposes that a blended feedback approach is adopted to maximize the students' learning outcomes.

***Keywords*:** Automated Writing Feedback, teacher feedback, writing performance, blended feedback approach

1. Introduction

Among major linguistic skills, writing is particularly challenging for learners to achieve mastery, often due to the lack of both practice opportunities and personalized feedback from educators (Hyland, 2007; Storch & Wigglesworth, 2010). During the writing process, formative feedback and opportunities for revision play a pivotal role in assisting learners in acquiring competence (Bardwell, 1981), not only by pinpointing errors but also serving as a powerful pedagogical tool for guidance (Perdana et al., 2025). Timely and thorough feedback provision, however, is particularly testing due to external factors including a high student-teacher ratio (Chen, 2025) and a heterogeneous distribution of proficiency levels within the same class (Alnemrat et al., 2025), together with the tutor's emotional fatigue due to overwhelming workload (Ballet & Kelchtermans, 2009). Consequently, this causes delays in feedback delivery and risks weakening instructional efficacy (Makwana, 2025).

Recent advances in technology, particularly in the field of automated writing evaluation (AWE), eg, ChatGPT, have offered promising solutions for the aforementioned issues. By providing real-time computerized feedback, ChatGPT addresses learners' concerns about various

aspects of their own writing in a timely manner, which in turn reduces the workload of teachers (Alnemrat et al., 2025). Multiple past studies have advocated a blended approach, in which the immediacy of automated feedback is strategically combined with nuanced and contextualized human feedback to achieve optimal correction efficacy (Makwana, 2025; Marzuki et al., 2023; Wu, 2024). To achieve this, it is crucial to understand the extent to which each feedback modality has a positive impact on learners' writing competence and their own perceptions of these feedback sources. Previous studies have often examined AI feedback from a general viewpoint, such as speaking, reading, accomplishment, or motivation (Mohammed & Khalid, 2025) in non-Vietnamese settings. Meanwhile, very few studies focus on the context of Vietnamese university settings, with Vietnamese students' beliefs and self-evaluation of the use of AI in writing correction remaining a fairly closed book. Furthermore, more research is also needed to explain what remains unknown about feedback quality, feedback uptake, and learner trust in AI-supported writing revision. To fill this gap, the purpose of this study is to examine how the two feedback modalities affect Vietnamese English-major freshmen's writing development as well as their affective responses. By shedding light on these important aspects, this study aims to contribute to the general understanding of the huge impact of AI on revolutionizing the current educational setting. Proposed below are the research questions:

1. To what extent do AI-generated feedback and teacher-provided feedback improve students' writing performance?

2. What are students' perceptions regarding the usefulness of each feedback modality?

3. What is the student's perceived motivational value of AI feedback, in contrast to teacher feedback?

2. Literature Review

2.1. Why does feedback matter?

The essential role of corrective feedback was based on theoretical frameworks of constructivist learning theory and sociocultural theory, which posit that feedback assists students in their process of comprehending and internalizing language, especially writing conventions (Hyland, 2007; Vygotsky, 1978).

Corrective feedback is a cornerstone in improving learners' writing performance, contributing to the promotion of long-term writing development towards greater proficiency (Makwana, 2025). Well-informed and formative feedback promotes learners' writing performance, provided it is delivered promptly, with a strong focus on meaning (Peltzer et al., 2024). First, through corrective feedback, learners can identify their own strengths and weaknesses to make informed revisions to their writing. In the realm of ESL writing, feedback clarifies erroneous aspects of the learner's written work, ranging from problematic linguistic forms (vocabulary, grammar, and mechanics) to rhetorical dimensions (organization, coherence, and argumentation) (Bitchener & Ferris, 2012). Secondly, feedback promotes learners' self-regulation and boosts their motivation. In the traditional classroom setting, teacher-provided feedback is particularly beneficial because students trust their teachers, who rely on their contextual knowledge to offer nuanced scaffolding in line with classroom dynamics (Alnemrat et al., 2025).

2.2. Traditional teacher-led instruction

Conventional correction practice, particularly teacher-led feedback, is central to ESL education. The multiple strengths of teacher-led feedback are evident in past studies. For instance, with their expertise, teachers can tailor their remarks to specific writing tasks to provide personalized and contextualized feedback that can account for the specific needs of learners. (Han & Li, 2024). Far from merely providing direct revision guidance, teachers provide metalinguistic instructions to foster students' critical thinking abilities by encouraging them to address higher-order writing concerns, such as structure and content development (Chen & Cheng, 2008). Despite its effectiveness, extensive teacher-led feedback falls short in providing immediate, extensive feedback, especially to teachers with limited time and resources (Makwana, 2025; Zhang & Hyland, 2018).

Written corrective feedback (WCF) refers to written information or codes given by teachers to address errors in learners' writing with the goal of leveraging their writing performance and facilitating second language acquisition. At its core, two levels of feedback include lower-order skills, which include sentence-level skills, such as spelling, punctuation, grammar, or capitalization; or higher-order skills that deal with idea elaboration, self-reflection, or writing organization (Wilson & Czik, 2016).

2.3. Computer-mediated feedback in writing

Investigation into automated written feedback has yielded heterogeneous outcomes (Baz & Aksoy, 2025; Goldstein, 2004; Makwana, 2025; Yu & Xie, 2025). When guided by rigorously designed prompts, AWF has the potential to bring positive impacts on enhancing students' writing performance. Through feedback, students achieve gains in multiple aspects of language learning, including speaking, learning accomplishment, self-regulation, and enjoyment of reading. Additionally, computer-based feedback is capable of improving the feedback quality while lessening the teacher's workload (Makwana, 2025; Yu & Xie, 2025) by generating immediate, specifically-tailored, and individualized feedback on the content and the language of students' writing (Lin & Crosthwaite, 2024), so that teachers can allocate more time to addressing higher-order concerns.

However, it should be noted that the effects of computerized feedback are not always positive. Past findings have indicated that this type of feedback tends to emphasize surface-level errors while overlooking meaning-level issues. Consequently, this does not contribute to broadening EFL learners' understanding of complex linguistic forms. Due to the nature of relying on algorithms to provide pre-defined errors and functions, these automated writing evaluation systems are unable to provide feedback that considers nuanced differences between learners. What is more, computer feedback is occasionally vague, lengthy, formulaic, or even misleading. Unfocused comprehensive feedback overloads the students' attentional capacity, which impedes the quality of writing and the overall writing progress.

2.4. ChatGPT as the source of feedback

Unlike past AWF platforms, ChatGPT, a product of GenAI, is capable of formulating human-like conversations through prompt construction (Han & Li, 2024). By definition, a prompt is a natural language text provided by the user to insert related information to inform the model to perform specific tasks, which allows the model to create output not originally part of its training

data (Si et al., 2022). Depending on the clarity and rigor of prompt formulation, ChatGPT offers various responses. Inheriting features from its predecessors, ChatGPT is also capable of giving instant formative feedback on different concerns of writing by clarifying major points (Makwana, 2025; Yu & Xie, 2025).

In summary, AI is capable of providing timely and repeated feedback, while teachers are able to provide contextual, motivational, and pedagogically sensitive feedback. By incorporating the strengths of these two feedback modalities, teachers can offer a more thorough feedback experience to their students, which advocates for a blended use of both feedback sources.

3. Methodology

3.1. Research participants and settings

Research participants include 54 Vietnamese English-major first-year students (21 males, 33 females, aged from 18 to 21). To minimize external variables, they were all taught by the same teacher (with a master's degree in Applied Linguistics and at least 10 years of experience in teaching writing for intermediate-level students) with the same writing lesson content. Based on the entrance exam result, their level was A2/B1 of the CEFR at the time the research was conducted. Most of them had prior experience in using AI-supported writing platforms such as ChatGPT and used these regularly on their personal digital devices, like smartphones or laptops. On the orientation day, after being clarified about the purpose of the research, these students were divided into two groups: 33 students received feedback from the teacher, and 21 students received it from ChatGPT. They were assigned based on their preferences and experience with using AI. Students could choose which group would best suit them, which led to unequal participant numbers. Therefore, this is a quasi-experimental comparative design rather than a randomized experiment. Therefore, to counteract the likelihood of variations in competencies between the two groups, an independent samples T-test was employed.

3.2. Research design and instruments

As quantitative writing scores and qualitative perception data were collected during the same intervention and interpreted together, the study adopted an embedded mixed-methods design. This research design was selected to provide a holistic view of not only what changes (writing scores), but also why and how they did so. In addition, numerical results on writing score gains alone will not suffice to state which feedback modality is more effective, as they do not provide insights into the cognitive and affective processes students undergo during interactions with feedback. Combining two sets of data is, therefore, crucial for a deeper understanding of the impact these two types of feedback have on students' writing performance.

Different research instruments were used to collect both quantitative and qualitative data. These include 8 paragraph writing topics taken directly from the textbook. The students are familiar with the major themes in the textbook (including daily activities, travel experiences, or an effective advertisement), and the topics are arranged to be progressively more challenging. The study employs a 5-point analytic rubric description table for assessment of 4 surface-level aspects (Grammar, Vocabulary, Mechanics, Sentence Structure) and 3 meaning-level aspects (Content and Idea development, Coherence and Organization, Writing style) (see Appendix B); a 5-point Likert-scale survey (Appendix C); and a set of questions for semi-structured interviews (Appendix D). This analytic rubric was adopted from various past studies (Alnemrat et al., 2025; Ferretti &

Graham, 2019; Qin & Karabacak, 2010), which helps analyze specific dimensions, rather than only giving a holistic impression of the student's writing ability. To ensure practicality, the rubric received reviews from the raters about whether the dimensions are suitable for paragraph writing evaluation. Based on the raters' expertise, the researchers made modifications to the rubric that best suited the purpose of the study. On the orientation day, 10 students were selected to write a paragraph as part of the pilot test. The raters marked these paragraphs, using the piloted rubric, to confirm its usability before officially being put into use for the remaining topics.

The content validity of the marking rubric, survey questions, and interview questions is supported. To begin, the marking rubric covers the seven key criteria of writing evaluation, divided into two main categories of micro issues (Grammar, Vocabulary, and Mechanics) and macro issues (Content and idea, Coherence and organization, and Writing style). Each criterion is classified into 5 levels, from poor performance (1) to excellent performance (5), to minimize the raters' bias and illustrate the aspects AI or teacher feedback benefits students most and least. Secondly, the survey confirms the construct validity by raising questions to explore the perception of students, including satisfaction, clarity, trust, and motivation. Additionally, the survey is divided into two versions, which compare two groups of students on the same dimension. The combination of quantitative data (Likert scale from questions 1-7) and qualitative data (question 9) allows for students' further explanation of their own choices, enriching data collection. Lastly, the interview questions dive deeper into affective factors of learners by mentioning important aspects such as "non-judgemental tool" and "anxiety". Also, questions regarding students' own critical thinking and learner agency (question 5) shed light on the learning process of students. Moreover, by asking students to compare the pros and cons of using AI in question 6, the interview encouraged them to reflect on their own learning process with AI, thus giving both the researcher and them a better understanding of the proper use of AI in improving their own writing quality.

3.3. Data collection

Data collection took place over 10 weeks, as described in the timeline below. During the orientation day (Week 1), after being informed of the purpose of the research, students provided their consent to participate. Topic 1 was also assigned, and students had one week to hand in their first version. Students in the AI group also received the assigned prompts for feedback from ChatGPT (see Appendix A). Well-crafted prompts minimize external variables due to inadequate prompt engineering and ensure that the AI responses are consistent among all students, which is essential in resulting in representative outcomes (Kung et al., 2023). To this end, the designed prompt only required AI to provide explicit suggestions for corrections, rather than correcting the whole paragraph itself. Students were advised to engage in continued dialogues with AI for deeper engagement with the feedback to fulfill the writing task.

After all students from both groups have submitted Draft 1, they revise it and submit the refined versions, called Draft 2. During the scoring, two raters, with years of experience in teaching writing, were trained on how to use the analytic rubric; any disputes over scoring were settled by their discussion for the final result. To further ensure objectivity, fairness, and consistency in writing evaluation, the Intraclass Correlation Coefficient (ICC) has been employed, with the results of 0.84 and 0.89 for pre-test and post-test, respectively.

Table 1

Intraclass Correlation Coefficient

Tests	Number of raters	ICC value	*p-value*	Reliability
Pretest	2	0.84	< .001	Good
Posttest	2	0.89	< .001	Good

In Week 10, all students took part in the survey with both closed and open-ended questions regarding their experiences and preferences for both types of feedback. 10 random students (5 from each group) were selected for an interview to gain deeper insights into their perspectives for a more comprehensive understanding of the efficacy of feedback.

Table 2

Description of the procedure

Timeline	Session description	Assignment of a new topic	1st version submission	2nd version submission
Week 1	Orientation + Instruction for Topic 1 (Pre-test)	Topic 1		
Week 2	Instruction for Topic 2	Topic 2	Topic 1	
Week 3	Instruction for Topic 3	Topic 3	Topic 2	Topic 1
Week 4	Instruction for Topic 4	Topic 4	Topic 3	Topic 2
Week 5	Instruction for Topic 5	Topic 5	Topic 4	Topic 3
Week 6	Instruction for Topic 6	Topic 6	Topic 5	Topic 4
Week 7	Instruction for Topic 7	Topic 7	Topic 6	Topic 5
Week 8	Instruction for Topic 8 (Post-test)	Topic 8	Topic 7	Topic 6
Week 9			Topic 8	Topic 7
Week 10				Topic 8

4. Results

To ensure baseline equivalence for the initial writing ability of the control group (teacher-led) and experimental group (AI-provided), the Shapiro-Wilk test was employed. This test was deliberately chosen to deal with samples with sizes less than 50 (Razali & Wah, 2011).

Table 3

Normality distribution of two groups

Test	Sample (N)	Shapiro-Wilk Statistic (W)	*p*-value	Normality Assumption

Teacher	33	0.963	**0.306**	Normal
AI	21	0.942	**0.237**	Normal

As the numbers between the two groups were unequal (33 and 21), the researcher also employed Levene's test to assess the equality of variances and confirmed that the *p*-value was greater than .05.

Table 4

Normality distribution of two groups

Variances	F-value	*p*-value	Equality of variances
Gain score	0.681	.412	Equal
Post-test Total	0.134	.715	Equal

Levene's test indicates that equality of variances was not violated (F = 0.681, *p* = .412). A paired samples t-test was run to evaluate the efficacy of feedback in writing improvement. Both groups achieved statistically significant improvement.

Table 5

Paired Samples T-test

Groups	**Test**	**Mean**	**SD**	***t***	***df***	***p (2-tailed)***	***Cohen's d***
Teacher (n=33)	Pre-test	19.82	4.31	-16.85	32	**< .001**	**2.93**
	Post-test	26.79	4.41				
AI (n=21)	Pre-test	19.14	3.52	-14.65	20	**< .001**	**3.20**
	Post-test	26.33	4.02				

The score gain of the AI group was slightly higher than that of the teacher group. The effect size was also calculated by using Cohen's d. At $d = 0.09$ (< 0.2), this shows that the difference in effectiveness between the teacher feedback and AI feedback is negligible, indicating that both groups improved during the intervention under different feedback conditions.

Table 6

Gain scores

Variance	***Group***	***Mean Gain***	***SD***	***t***	***df***	***p (2-tailed)***	***Cohen's d***
Gain scores	*Teacher*	*6.97*	*2.38*	*-0.325*	*52*	***.746***	*0.09 (< 0.2)*
	AI	*7.19*	*2.25*				

Criteria	***Gain score (Teacher)***	***Gain score (AI)***	***p-value***	***Result***

Grammar	*1.36*	*1.16*	*0.452*	*Similar*
Vocabulary	*0.88*	*0.72*	*0.485*	*Similar*
Mechanics	*0.88*	*1.13*	*0.354*	*Similar*
Sentence Structure	*1.06*	*0.88*	*0.428*	*Similar*
Content & Idea	*0.94*	*1.03*	*0.697*	*Similar*
Coherence & Organization	*0.97*	*1.00*	*0.895*	*Similar*
Writing Style	*0.88*	*0.94*	*0.812*	*Similar*

Table 7

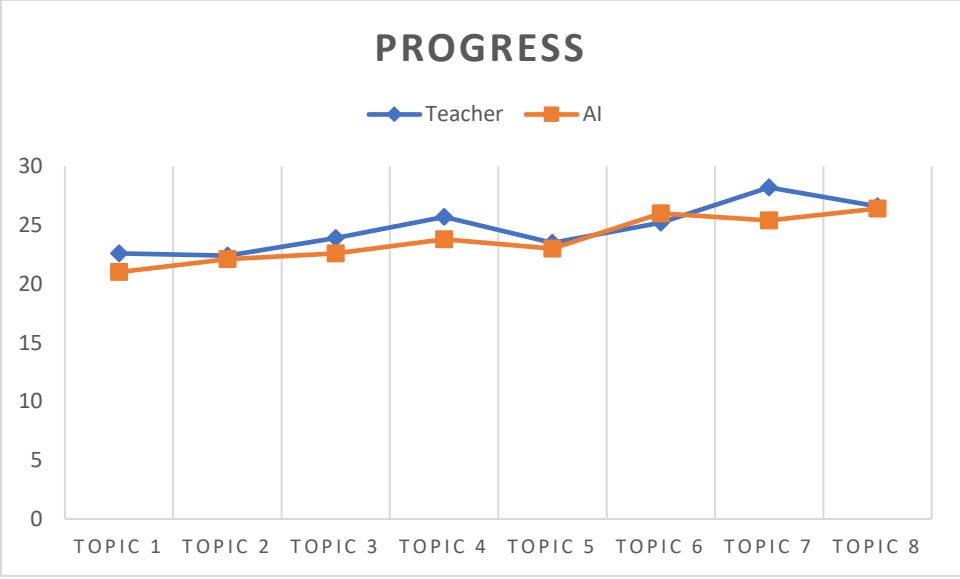
Mean scores across topics of both groups

Group	***Topic 1***	***Topic 2***	***Topic 3***	***Topic 4***	***Topic 5***	***Topic 6***	***Topic 7***	***Topic 8***
Teacher	22.6	22.4	23.9	25.7	23.5	25.2	28.2	26.6
AI	21.0	22.1	22.6	23.8	23.0	26.0	25.4	26.4

The mean scores of the rewritten versions for the eight topics were recorded. These versions were the refined products of the students after receiving either of the two feedback types.

Figure 1

Progress in scores of two groups for 8 weeks



For the first four weeks, both groups enjoyed fairly gradual improvement in their writing performance, with the figures climbing from 22.6 to 25.7 for the teacher group, and 21 to 23.8 for the AI group. During this period, the teacher group slightly outperformed the AI group. In week

5, the scores of both groups fell, which was possibly due to the difficulty level of this week's topic. With both groups falling in performance, this signifies that the difficulty level was consistent for both. During the remaining period, it is notable that even though the teacher group retained its constant growth, the AI group improved significantly and outperformed its counterpart for the first time in week 7 (28.2 compared to 25.4), before reaching the same score level in week 8 at just over 26.0. This implies that the AI group might require more time to become acquainted with the novel feedback method.

Survey responses

The reliability of two questionnaires (for the teacher group and the AI group) was confirmed with Cronbach's alpha. The figures for both were higher than 0.8, indicating high reliability. All items had a total correlation of at least 0.30; thus, none were removed from the questionnaire. To sum up, all items were relevant and valid to collect insights from participants regarding multiple aspects of their writing experience, including satisfaction levels, trust in feedback accuracy, perceived usefulness to both surface- and meaning-level issues, perceived motivation, and overall writing performance improvement.

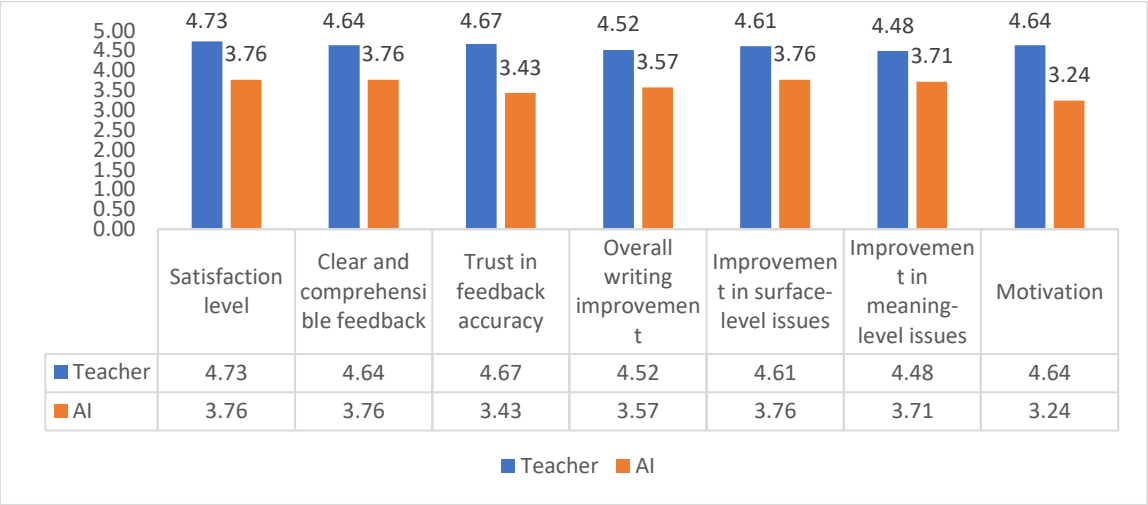
Table 8

Reliability of the questionnaire

Surveyed groups	Sample (N)	Number of items	Cronbach's Alpha	Evaluation of reliability
Teacher	33	7	**0.932**	**Reliable**
AI	21	7	**0.871**	**Reliable**

Figure 2

Survey responses of Teacher group and AI group

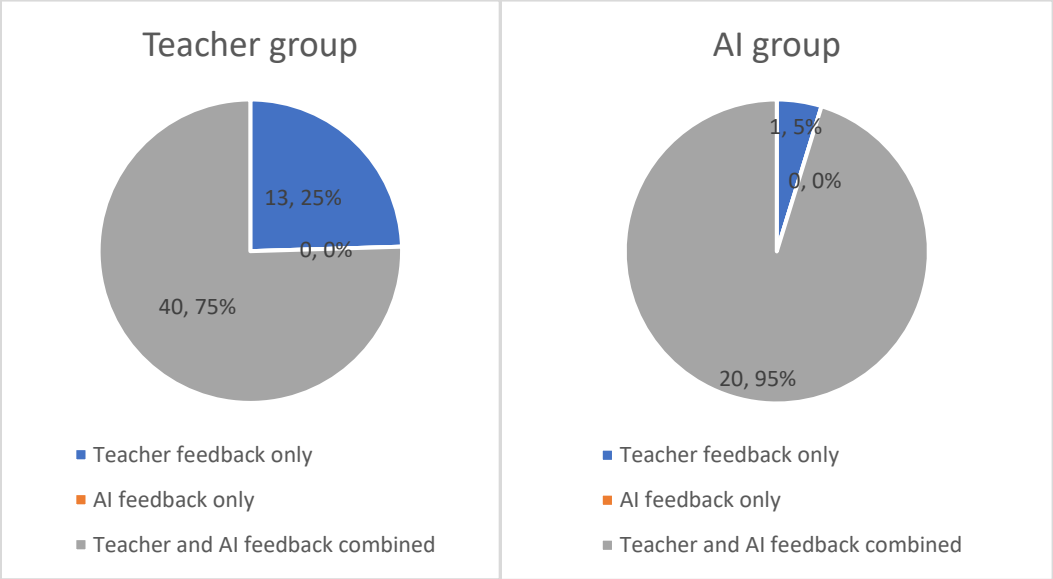


Overall, compared to AI-generated feedback, teacher-provided feedback was better received in all categories. The aspects that students appreciated most were their satisfaction level with the teacher-provided feedback (4.73) and their trust in their instructor's ability to deliver accurate corrections (4.67). Meanwhile, the aspects that students found most content with using AI were the general satisfaction level, clarity and comprehensibility of the feedback, and

improvement in surface-level issues, at 3.76 each. Similar gain scores in this sample suggested that the aspect in which AI was perceived to be nearly as useful as a human teacher was meaning-level issues (content development and organization), with a difference of just 0.77. By contrast, the widest gap was seen in the motivation aspect, with a difference of 1.4, indicating that AI was less effective in motivating students to revise their own writing and improve their later writing assignments.

Figure 3

Preferences for feedback sources



Responses from the survey highlight a common pattern in students' perceptions of feedback sources: a strong preference for a blended approach of both feedback modalities, and an absence of fondness for AI feedback alone. Within the teacher feedback group, a quarter of students expressed a liking for teacher feedback solely, while this figure was three times larger (75%) in those who opted for a combination of both feedback types. Similarly, a vast majority of students from the AI-feedback group (95%) preferred this dual feedback, with only one student who favored the traditional feedback type. No student from either group placed their full trust in AI alone.

To dive even deeper into the learners' perspectives on the feedback sources, 5 students from each group were invited to join the interview (with a total of 10 students). With their consent, the whole interview was recorded for later transcription and analysis. A thematic approach was employed to classify the interview responses into 12 major themes, which helped shed light on the pros and cons of each feedback type. Irrelevant or vague responses, such as "OK", were left out. The 12 major themes were collected from responses to interview questions, then generated by the following procedure. First, the interview recordings were transcribed into text. The researcher read these scripts multiple times to understand the context and significance and identify key features. Second, these features are grouped into categories with similar features. For instance, codes such as "quick correction" and "didn't have to wait" were grouped into "AI provides quick responses".

Thirdly, the list of themes is narrowed down, with only themes that are mentioned repeatedly selected. Illustrated below are the 12 major themes (6 for each group) and the quotes from the interviewees. The number of students who mentioned these themes was also included.

Table 9

Students' personal perspective on Teacher feedback

Major themes	**Quotes**	**Number of students who mentioned this**
1. Teacher provides detailed corrections	"…teacher's correction is detailed, helping me understand better and improve my skills"	3/5
2. Teacher provides comprehensible feedback	"…teacher's correction is very easy to understand"	3/5
3. Teacher's professionalism allows highly accurate and trustworthy feedback	"…teacher understands my writing problems and corrects my writing in line with the writing task"	4/5
4. The teacher understands the students' context, thus providing feedback that aligns with the writing goal	"…teacher gave me highly professional comments that suit the requirements of the lesson"	3/5
5. Teachers help develop meaning-level aspects (idea development)	"…teacher provides feedback to help me extend the ideas for my writing"	3/5
6. Teacher's "human touch" boosts the student's writing motivation and helps them feel at ease	"…sometimes my teacher adds jokes and interactions, which makes me feel interested when receiving feedback"	5/5

Table 10

Students' personal perspective on AI feedback

Major themes	**Quotes**	**Number of students who mentioned this**
7. AI provides quick responses	"…feedback is provided constantly as I ask questions or make requests, so I didn't have to wait"	5/5
8. AI feedback is vague and mechanical, lacking depth	"…sometimes AI feedback is unclear to me, and it lacks depth"	3/5
9. AI feedback is convenient, allowing students to engage in prolonged conversations for clarifications	"AI can support me anytime, so I can study whenever I need to"	5/5
10. AI excels at spotting and correcting surface-level errors	"AI might do better in correcting minor errors that teachers might overlook"	2/5
11. AI still occasionally makes mistakes	"AI can point out the mistakes with grammar and vocabulary, but not entirely correct all the time"	2/5

12. AI's lack of human emotions might not motivate students to revise their writing	"(AI's) disadvantage: lack of emotions, not as motivating as teachers"	3/5

5. Discussion

1. To what extent do AI-generated feedback and teacher-provided feedback improve students' writing performance?

As indicated by the results, both feedback types had a positive impact on students' writing performance, with the teacher group's score rising from 19.82 to 26.79 and the AI group from 19.14 to 26.33. In terms of the difference in effectiveness between each feedback type, the tests suggested that both teachers and AI were helpful in leveraging students' writing ability. A detailed look into the progress reveals that in the first half of the period (week 1 – 4), the teacher group slightly outperformed the AI group. Presumably, this could have been the degree of familiarity of students with the feedback modality. While the teacher group was familiar with the conventional feedback strategy and applied it more effectively (their score increasing from 22.6 to 25.7), it took longer for the AI group to become accustomed to using AI for writing revision (with their score increasing from 21.0 to 23.8). After this threshold, however, the AI could keep up with the teacher group in the following weeks, even surpassing them in week 6 (26.0 compared to 25.2). In the final week, both groups reached the same score at approximately 26.

Regarding feedback levels, past studies argued that while AI excelled at addressing local concerns, such as syntax, lexical resource, or spelling, teachers are superior in global concerns, including ideation or organization. In line with this, findings of this study also found that students in the AI group achieved higher scores for mechanics (including errors with spelling, punctuation, and capitalization), with 1.13 compared to 0.88. This is evident as AI's algorithmic nature allows it to compile information from multiple sources and formulate common rules. However, the same did not apply to grammar, as the grammar score of the teacher group was slightly higher than that of the AI group, with 1.36 compared to 1.16. It could be suggested that when teachers do attend to providing detailed corrections to surface-level errors, including grammar, they can also provide accurate feedback that allows students to clearly understand grammar rules.

With respect to meaning-level aspects, including idea development and organization, it is notable that within this specific sample, task type, prompt design, and short intervention period, AI feedback was just as effective as teacher feedback (1.03 and 1.00 compared to 0.94 and 0.97, respectively). This contradicts past studies, which claimed that the realm of higher-order thinking skills was the strength of human teachers. This could be due partly to recent advancements in AI technology, whose database has been constantly growing. Arguably, careful prompt engineering might also have a role to play, as the prompt requires AI to also develop students' logical thinking, which helps produce not only accurate but also convincing pieces of writing.

2. What are students' perceptions regarding the usefulness of each feedback modality?

Students in the sample value teacher feedback in all seven aspects, namely overall satisfaction, clear and comprehensible feedback, trust in feedback accuracy, overall writing improvement, improvement in surface-level issues, improvement in meaning-level issues, and

motivation. Interestingly, this reveals that although the students benefit from AI feedback and enjoyed score gains, they still gravitate towards teacher feedback. One reason for this could be that students are still not totally familiar with interactions with AI and thus might develop less trust for this new feedback modality, compared to teacher feedback. A further explanation for this strong preference could lie in the last aspect of the survey: motivation. This is the aspect with the greatest deviation in both groups. While the score for motivation of the teacher group was 4.64, that of the AI group was far lower, at 3.24. This recurring theme was reiterated in the survey's last question. With his understanding of the students' context, including their language proficiency, the teacher was able to provide feedback that not only matches the students' level but also aligns with the writing goals. The result of this was focused, tailored, comprehensible feedback that students could instantly understand and apply accordingly. Students' comfort and great trust in the expertise of their teacher were other reasons why they were more motivated to improve their writing in the following topics, compared to mechanical and vague feedback from ChatGPT, which was reported to be inaccurate at times, causing a sense of distrust. Participants in the interview complained that:

"AI didn't give me detailed corrections, nor feedback; it just generally states something like 'you made quite a few mistakes.'"

Despite this, students do appreciate the immediacy of AI, stating that the feedback was delivered in a timely manner upon request. This was helpful as the process of correction could start right after the students finish their writing, which avoids delays that might hinder learning. One student said:

"If the feedback arrives late, I might forget the ideas I put in my writing."

A further upside of AI is its availability, which allows users to ask questions anytime they are pleased, and the conversation can also be extended until the users are satisfied with the responses and their final writing product. Immediacy and availability have long been seen as shortcomings of teacher feedback, as teachers are usually overwhelmed with workload, resulting in delayed feedback, which might fail to provide students with the necessary information to refine their writing in time.

Responses from the interview highlight that students of both groups are in favor of a combined approach of both feedback sources. Those from the teacher-feedback group said that:

"I think we can use both and learn the good things from each of them."

"We can use AI, but we shouldn't completely depend on it, because good writing must contain human emotions."

3. What is the student's perceived motivational value of AI feedback, in contrast to teacher feedback?

Qualitative data collected from the class survey and the interview assist in measuring the students' motivation. Through semi-structured interviews, students express their thoughts on how the feedback modalities encourage them to put in the extra effort in revising their own writing. From the findings of this paper, students receiving teacher feedback confirmed a higher level of motivation compared to those receiving AI feedback (themes 6 and 12). This was in line with previous studies, which also claimed that the human factor was of greater value in boosting the learners' learning motivation (Alnemrat et al., 2025; Mohammed & Khalid, 2025). That said, the

effects of teacher feedback and AI feedback on learners' motivation in writing revision remain a multifaceted issue. On the one hand, students highly value the instant and easily accessible feedback AI generates, as this eliminates the long delays of feedback often associated with conventional teacher feedback. Without fear of forgetting what they have written, students can immediately ask AI to provide feedback and refine their writing, which gives them the motivation boost to complete their writing tasks quickly.

On the other hand, the human characteristics of teacher AI were a major contributor to students' motivation. Firstly, students' trust in the teacher's ability to provide accurate and well-suited feedback motivates them to improve their writing skills during revision. Compared to AI's generative responses, teacher feedback was more focused and tailored to students' comprehension. Previous studies have strongly supported this, as it avoids overwhelming students with redundant corrections on all aspects of their writing. Additionally, teacher feedback included words of encouragement and even small jokes, which significantly put students at ease and inspired them to write better. Interestingly, some students claim that when receiving the handwritten feedback of their teacher, they feel appreciative of their teacher's dedication to helping students and determined to correct their writing accordingly.

"I feel more motivated to think that my effort is recognized by my teacher, which encourages me to perform better as a way to repay my teacher's dedication."

Furthermore, it is worth considering the novelty effects and familiarity when discussing the impact on the motivational value of teacher feedback and AI feedback. Regarding the novelty effects, students tend to feel a great sense of enthusiasm during the first stage of the learning process as they experience a novel learning method. Despite short-term improvement in motivation, students may "blindly" accept AI quick corrections instead of truly learning, which may lead to loss of learner agency or over-reliance on AI tools, as warned by previous researchers (Shi et al., 2026). Familiarity may also come into play in affecting the learning outcomes. Students who have been too accustomed to the longstanding habits of receiving feedback from the educators might suffer from temporary anxiety and discomfort due to a change in unfamiliar learning methods (Bakla, 2020). Thus, the successful integration of modern learning tools heavily depends on whether the students are willing to abandon old learning habits. This will prove challenging for students who tend to trust their teachers' expertise and authority more than technological tools (Chen, 2025). The lack of contact with AI may reduce trust in this tool, as students may argue that AI lacks the authority to evaluate their writing (Yu & Xie, 2025). To sum up, by understanding the novelty effects of using AI, thus maintaining stable engagement with the technology and developing familiarity with the novel learning experience over the traditional method, students can gain a better understanding of the proper use of AWE and truly take advantage of this innovative learning approach.

6. Conclusion

Findings from the 10-week research have offered deeper insights into the current knowledge of the use of AI in providing feedback to improve students' writing skills. All students enrolled in the same course, used the same textbook, and received instructions from the same teacher, with the only difference being their feedback source. Relevant tests were applied to inform the researcher of the extent to which each feedback source benefited the students, and survey

responses were collected to add deeper meaning to the data by elaborating score gains and explaining students' perceptions.

In terms of the effectiveness of each feedback modality on students' writing improvement, it might be proposed that in this specific sample, task type, and short intervention period, AI feedback was nearly as effective as teacher feedback, provided that the prompt is carefully engineered and students took an active role in interacting with AI for deeper engagement with the feedback. Apart from the overall positive effect on writing gains, AI was also remarkably helpful in addressing various aspects of students' writing, including both surface-level errors and even meaning-level errors. Results of the findings stand to argue that with a well-tailored prompt and active engagement from students, AI is fully capable of providing in-depth suggestions on higher-order concerns of writing and propelling students into more proficient writers. However, it should be noted that AI does not always outperform teachers regarding surface-level issues such as grammar or spelling. This means that the argument relating to AI's better ability to correct surface-level issues might actually stem from the lack of teachers' time to extensively provide corrections for this type of issue. The implication of this is that the current AI technology is capable of assisting teachers in feedback provision, in the face of heavy workload and time constraints to mark large numbers of assignments. However, students are still in great need of their teacher's guidance on designing prompts that align with the course's objectives and providing them with professional and personalized comments that tie both feedback types together. Only at this stage will students reap the full benefits of corrective written feedback and make steady progress to become more proficient writers.

With respect to motivation, there will still be a long way to go for AI to provide the inspiring feedback that fully captures the students' interest and attention. This might stem from the students' attitude towards the role of AI in their own learning. Students tend to regard AI as a technological tool that offers instant and widely available suggestions, and although they value its quickness and vast knowledge, they do not fully trust it to provide accurate information, which might lessen the credibility of its feedback. By contrast, students have much greater confidence in their tutor, believing that the teacher's expertise will safeguard the accuracy of the feedback. This great degree of trust is the primary reason why students seem to take teacher feedback more seriously and are more likely to associate their improvement with their teacher's hard work and dedication.

Some students expressed concerns about information overload from AI feedback. To tackle this, it is advised that students be trained not only in how to devise proper prompts but also in how to select only relevant and useful computerized suggestions that benefit their own writing, instead of trusting all feedback generated by AI. This process will further sharpen their critical thinking skills.

This study acknowledges the following limitations. The first shortcoming is the small sample size (54) and differences in the number of students in the two groups (33 and 21). Despite relevant tests to ensure reliability, this might still impact the generalizability of the sample due to self-selection bias; thus, the findings should be seen as exploratory. Secondly, 10 weeks was a satisfactory period for investigation into the effectiveness of two types of feedback, but this cannot confirm whether the effect is long-lasting after the intervention came to an end. During this period, there was limited control over students' actual engagement with AI feedback outside class, so

external variables have not been fully regulated. Thirdly, the student's English proficiency level is intermediate, and their main task was to write in one paragraph, so there is a need for further research into other proficiency levels and with other types of writing tasks, such as creative writing. Lastly, even though past studies stated that teachers were less effective than AI in providing surface-level corrections, it is unclear whether teachers' willingness to provide detailed feedback was at play, together with some other variables, such as novelty effects and familiarity.

To address these shortcomings, the researcher therefore proposes that future studies can be carried out on larger sample sizes, with a wider variety of language proficiency, for more generalizability. Furthermore, a blended approach could be adopted to gauge the effectiveness of both feedback types. This means instead of receiving feedback from AI alone, the experimental group could receive feedback from both AI and the teacher to test if there would be a cumulative effect that far surpasses a single type of feedback. Future researchers could also examine feedback uptake directly by analyzing the process by which students revise from Draft 1 to Draft 2. This would provide stronger evidence of how teacher and AI feedback actually affect revision behavior, rather than relying solely on score gains and perceptions.

To conclude, this research attempts to contribute to answering the rising question, "Can AI replace teachers?". It can be argued that AI's positive effect in EFL education should not be underestimated, but students will still need guidance from human tutors to utilize this tool efficiently and responsibly. What is more, the affective factors, such as trust and motivation, shall continue to secure the teacher's role in the current climate of education in the information era. However, in order to truly leverage their pedagogical capabilities and better adapt to the ever-changing waves of technology, teachers are advised to take full advantage of AI's power to provide their learners with a novel learning experience that not only enhances their linguistic skills but also facilitates their lifelong learning.

References

Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education, 10*. https://doi.org/10.3389/feduc.2025.1614673

Bakla, A. (2020). A mixed methods study of feedback modes in L2 writing. *Language Learning & Technology, 24*(1), 143–161. https://doi.org/10.125/71192

Ballet, K., & Kelchtermans, G. (2009). Struggling with workload: Primary teachers' experience of intensification. *Teaching and Teacher Education, 25*(8), 1150-1157. Retrieved from https://doi.org/10.1016/j.tate.2009.02.012

Bardwell, R. (1981). Feedback: How does it function? *Journal of Experimental Education*, *50*(1), 4–9. https://doi.org/10.1080/00220973.1981.11011792

Baz, M. A., & Aksoy, S. H. (2025). The effect of feedback on informative text writing: AI or teacher? *Open Praxis, 17*(3), 611–631.

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing.* Routledge.

Chen, C.-F. E., & Cheng, W.-Y. E. C. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology, 12*(2), 94–112.

Chen, Q. (2026). Students' perceptions of AI-powered feedback in English writing: Benefits and challenges in higher education. *International Journal of Changes in Education, 3*(3), 305-315. https://doi.org/10.47852/bonviewijce52025580

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: theory, assessment, and instruction. *Reading and Writing*. https://doi.org/10.1007/s11145-019-09950-x

Goldstein, L. M. (2004). Questions and answers about teacher written commentary and student revision: Teachers and students working together. *Journal of Second Language Writing, 13*(1), 63–80. https://doi.org/10.1016/j.jslw.2004.04.006

Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System, 126*, Article 103502. https://doi.org/10.1016/j.system.2024.103502

Hyland, K. (2007). Genre pedagogy: language, literacy and L2 writing instruction. *Journal of Second Language Writing.* https://doi.org/10.1016/j.jslw.2007.07.005

Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System, 127*, Article 103529. https://doi.org/10.1016/j.system.2024.103529

Makwana, V. (2025). A comparative analysis of AI-powered and teacher-led feedback: Investigating student perceptions and writing performance. *Journal of English Language Teaching, 67*(1), 3–12.

Marzuki, Widiati, U., Rusdin, D., Darwin, D., & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2). https://doi.org/10.1080/2331186x.2023.2236469

Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: Mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia*. https://doi.org/10.1186/s40468-025-00343-2

Peltzer, K., Lorca, A. L., Krause, U.-M., & Busse, V. (2024). Effects of formative feedback on argumentative writing in English and cross-linguistic transfer to German. *Learning and Instruction, 92*, Article 101935. https://doi.org/10.1016/j.learninstruc.2024.101935

Perdana, Indra & Bungai, Joni & Ilham, Ilham. (2025). The Students' Perceptions of AI-Generated and Teacher Feedback in English Writing Skills: A Comparative Study. *Jurnal Pendidikan Bahasa Inggris Undiksha. 13*. 22-31. 10.23887/jpbi.v13i1.92668.

Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System, 38*(3), 444–456. https://doi.org/10.1016/j.system.2010.06.012

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21–33.

Shi, H., Chai, C. S., Zhou, S., & Aubrey, S. (2025). Comparing the effects of ChatGPT and automated writing evaluation on students' writing and ideal L2 writing self. *Computer Assisted Language Learning, 39*(4), 1148–1175. https://doi.org/10.1080/09588221.2025.2454541

Si, C., Gan, Z., Yang, Z., Wang, S., Wang, J., Boyd-Graber, J., & Wang, L. (2022). *Prompting GPT-3 to be reliable* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2210.09150

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, *32*(2), 303–334. https://doi.org/10.1017/S0272263109990532

TH, K., M, C., A, M., C, S., L, D. L., C, E., M, M., R, A., G, D.-C., J, M., & V., T. (2023). Performance of chatgpt on USMLE: potential for ai-assisted medical education using large language models. *PLOS Digit Health, 2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers & Education, 100*, 94–109. https://doi.org/10.1016/j.compedu.2016.05.004

Wu, L. (2024). AI-based Writing Tools: Empowering Students to Achieve Writing Success. *Advances in Educational Technology and Psychology*, *8*(2). https://doi.org/10.23977/aetp.2024.080206

Yu, H., & Xie, Q. (2025). Generative AI vs. teachers: Feedback quality, feedback uptake, and revision. *Language Teaching Research Quarterly, 47*, 113–137.

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

Bionote

The author is currently working in tertiary education at the English Faculty. He has been teaching English for different purposes (communication, test preparation) for over 10 years. During his career, he has developed a fascination with the dynamics and rapport between teachers and learners. With a growing curiosity about how this interpersonal relationship facilitates the learning process and outcome, he has sought to pinpoint aspects that will render human educators irreplaceable, especially in light of the prevalence of AI technology.

Appendices

Appendix A

Prompt: *You are my writing mentor. I will paste my paragraph draft below, based on the paragraph structure (topic sentence – supporting sentences – concluding sentence). Please give me feedback without rewriting the paragraph.*

Follow these steps:

A. Start with strengths: Tell me what I did well in my paragraph.

B. Then, give me at least one suggestion for improvement for each aspect below:

1. Grammar (word form, article, tense)

2. Vocabulary / Word Choice (is the choice appropriate and natural?)

3. Mechanics (spelling, punctuation, capitalization)

4. Sentence Structure (variety and complexity)

5. Content and Idea Development (is the idea well-developed, with main points and supporting ideas?)

6. Coherence and Organization (are the ideas well-connected, with good use of linking words?)

7. Writing Style (is the style appropriate and natural?)

C. Give an overall comment: Summarize the main things I should focus on when revising.

D. Remember: Do not rewrite my paragraph — just guide me with feedback.

Appendix B

Marking rubric

Criteria	1 (very weak)	2 (weak)	3 (average)	4 (good)	5 (excellent)	re
mmar	quent grammar errors make it very icult for the reader to understand.	quent grammar errors, but text can still be understood.	e grammar errors, but they do seriously affect clarity.	grammar errors; uses ence structures relatively ibly and accurately.	ng command of grammar; flexible and urate use of sentence structures. Little to rrors.	
abulary / Word oice	y limited vocabulary: frequent rrect word uses causes understanding.	ited vocabulary, repetitive d choice. Some propriate word usage.	s vocabulary at a basic level, ropriate to the topic.	s a variety of appropriate abulary, making the writing ly.	n vocabulary; precise and sophisticated d choice showing deep understanding of ning.	
chanics (spelling, ctuation, italization)	ny spelling and punctuation errors ear frequently, distracting the ler.	ny spelling and punctuation rs reduce the quality of the ing.	e spelling or minor punctuation rs, not significant.	spelling and punctuation rs; used effectively.	lling, punctuation, and capitalization are pletely accurate and careful.	
tence Structure riety & complexity)	ny spelling and punctuation errors ear frequently, distracting the ler.	tences have simple ctures with little variety.	ws connection between ences and paragraphs, but etimes awkward or repetitive.	s a variety of sentence ctures (simple, compound, plex) relatively effectively.	tence structures are used flexibly, urately, and effectively, creating a oth flow.	
ntent & Idea elopment	clear main idea or ideas are vague. evidence to support arguments.	n ideas are weak; supporting s are not persuasive enough.	n idea is clear but supporting s are still limited or not fully eloped.	n ideas are clear; supporting s are well developed with cific evidence.	y clear main idea: supporting ideas are iled, logical, and insightful, onstrating strong critical thinking.	
erence & anization	tences and ideas are scattered, ing connection.	organized arrangement; icult for readers to follow.	s are logically arranged, but sitions between sentences and graphs are not smooth.	tences are logically arranged; ctive transitions create erence.	entire piece flows smoothly; ideas are tly connected and presented naturally.	
iting Style	ting style is unclear, unnatural, or s coherence, making it difficult for lers.	ting style is inconsistent, several inappropriate ressions.	ting style is relatively ropriate to the topic but not very nt.	ar, confident writing style ropriate to the topic and genre.	ural, engaging writing style reflects the er's voice appropriately.	

TOTAL (7-35)

Appendix C

SURVEY

Part 1 - Participants' details:

1. Full name: _____
2. Gender: ☐ Male ☐ Female
3. ID (written by the researcher): _____

Part 2 – Survey questions:

Construct	***AI-feedback group***	***Teacher-feedback group***
Satisfaction level	1. I am **satisfied** with the feedback I received from AI throughout the 10-week writing course.	1. I am **satisfied** with the feedback I received from the teacher throughout the 10-week writing course.
Clarity and comprehensibility	2. AI feedback is **very clear and easy to understand**, helping me know exactly what I need to do to improve my writing.	2. The teacher's feedback is **very clear and easy to understand**, helping me know exactly what I need to do to improve my writing.
Trust in feedback accuracy	3. I **trust** the **accuracy** of AI's revision suggestions.	3. I **trust** the **accuracy** of the teacher's revision suggestions.
Overall writing improvement	4. AI feedback has helped me effectively improve my **overall writing skills**.	4. The teacher's feedback has helped me effectively improve my **overall writing skills**.
Improvement in surface-level issues	5. AI feedback helps me improve **grammar, vocabulary, punctuation, capitalization, and spelling** (surface-level errors).	5. The teacher's feedback helps me improve **grammar, vocabulary, punctuation, capitalization, and spelling** (surface-level errors).
Improvement in meaning-level issues	6. AI feedback helps me improve **writing organization, idea development and argumentation, and the logical sequencing of ideas within sentences** (meaning-level errors).	The teacher's feedback helps me improve **writing organization, idea development and argumentation, and the logical sequencing of ideas within sentences** (meaning-level errors).
Perceived motivation	7. AI feedback gives me **motivation** to **take the initiative to revise and write better** in subsequent assignments.	The teacher's feedback gives me **motivation** to **take the initiative to revise and write better** in subsequent assignments.
Preferences for a blended approach	8. If given the choice, would you prefer to receive feedback from the teacher, from AI, or a combination of both?	

Personal insights	9. Please explain your choice above in detail and describe the greatest advantage and disadvantage of the type of feedback you received (teacher or AI).

Notes:

- From questions 1-7, participants choose 1-5 points on the Likert scale, of which:

1: Totally disagree

2: Disagree

3: Neutral

4. Agree

5. Totally agree

- In question 8, participants choose between three options:

☐ Teacher feedback only ☐ AI feedback only ☐ Teacher and AI feedback combined

- Question 9 is an open-ended question to explore aspects that may not have been covered in the previous questions.

Appendix D

INTERVIEW QUESTIONS (AI-FEEDBACK GROUP)

1. How would you evaluate the level of detail in the AI's feedback compared with what you would expect from a teacher? Do you think the AI was more effective at helping you improve surface-level errors (e.g., grammar and spelling) or meaning-level issues (e.g., organization and idea development)?

2. What benefits did receiving almost immediate feedback from the AI bring to your learning process? Did you find the AI's feedback to be consistent across different writing submissions?

3. How did receiving feedback from a non-judgmental tool like AI affect your writing? Did it make you feel more confident or less anxious when submitting your work?

4. Did you encounter any difficulties when reading and applying the AI's feedback? For example, did the AI ever use vocabulary or expressions that you had to look up? If so, did this make it harder for you to understand and apply the feedback?

5. Did you ever feel that the AI's feedback was repetitive or suggested changes that you thought were unnecessary? Did you feel the need to question or respond to the AI's feedback before accepting its suggestions?

6. Compared with the traditional feedback you have previously received from teachers, in what ways do you think AI feedback is superior (e.g., speed, specificity, comprehensiveness)? Conversely, what strengths does teacher feedback have that AI cannot replace?

7. Would you recommend that other students use AI-generated feedback as a self-learning tool? In your opinion, at which stage of the writing process should AI be used most effectively (e.g., brainstorming ideas, producing a first draft, or making final revisions)?

#118 STUDENT PERCEPTIONS OF CHATGPT-GENERATED FEEDBACK IN IELTS SPEAKING PART 3 PRACTICE: A QUALITATIVE STUDY

Diem Bich Huyen Bui[1], Vo Man Nghi Nguyen [2]

Abstract

The integration of Artificial Intelligence (AI) tools in language learning has sparked interest in their potential use in supporting the improvement of speaking skills, especially in high-stakes tests such as IELTS. However, research about how learners perceive the usefulness of AI-generated feedback in IELTS Speaking Part 3 practice remains limited. Therefore, this study employed a qualitative approach to explore students' perceptions of using ChatGPT for IELTS Speaking Part 3 practice. Semi-structured interviews were conducted with eight English-major undergraduates with IELTS band scores ranging from 5.5 to 6.5 who had experience using ChatGPT to practice IELTS speaking part 3 and receive feedback. Both positive and negative perceptions were revealed using qualitative data analysis. On the one hand, ChatGPT was perceived as useful in providing a secure environment to practice speaking, giving immediate feedback, expanding lexical and grammatical resources, and developing ideas, coherence and cohesion of the speech. On the other hand, participants reported drawbacks in speech recognition and prompt quality. Such limitations were found to influence learners' trust in the system and shape their overall perception of its usefulness. The findings indicate potential applications of ChatGPT in facilitating IELTS Speaking Part 3 practice and, at the same time, highlight the need to provide training or guidance on creating efficient prompts and follow-ups, as well as to raise learners' awareness of critical evaluation of AI-generated responses.

***Keywords*:** AI-generated feedback, ChatGPT, IELTS Speaking Part 3, qualitative study

1. Introduction

In the era of globalization, excellent English skills are widely seen as a crucial measure of competitiveness (Wu et al., 2020). Tertiary educational institutions worldwide have placed emphasis on students' mastery of English as one of the requirements for academic pursuits (Muluk et al., 2022). As a result, demand for international English proficiency certificates has been on the rise, with the International English Language Testing System (IELTS) being among the most common tests nowadays. IELTS is one of the leading assessments in Vietnam as it is considered a primary criterion in many university admissions, graduate prerequisites, scholarship hunts for studying abroad, and career development (H. N. Nguyen & D. K. Nguyen, 2022). Therefore, achieving a high IELTS score can help individuals thrive in today's competitive educational context.

Vietnamese IELTS test-takers are often reported to perform better at receptive skills and struggle with productive skills. In the 2023-2024 period, the average band scores for Vietnamese

[1]*Corresponding author, PhD Candidate, School of Languages, International University – Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam; Email: bdbhuyen@hcmiu.edu.vn*
[2] *BA, Department, School of Languages, International University – Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam*

IELTS candidates were 6.3 in Listening, 6.4 in Reading, 6.0 in Writing, and 5.7 in Speaking (Nga, 2024). Speaking was particularly lower than the global mean and has been considered the most difficult skill for many EFL learners in Vietnam (Le, 2017).

Le (2017) also argued that, since the speaking test requires face-to-face interaction and spontaneous, direct answers to questions, candidates' anxiety and underperformance during this productive skill test are serious concerns. Vietnamese test-takers often struggle to develop ideas, maintain coherence and fluency, and produce immediate, relevant answers under pressure (Le, 2017). Traditional face-to-face methods often fail to address these issues for several reasons. First, traditional classes often focus on language aspects rather than skills development. Second, influenced by culture, many students have a habit of memorizing examples by heart rather than brainstorming ideas or generating their own responses (Duong & Nguyen, 2006). Finally, students lack exposure to authentic practice and immediate feedback due to large class sizes or a tight syllabus (Copland et al., 2014). As a result, many test-takers are still unable to achieve their target band scores despite intensive preparation.

The IELTS speaking test comprises three parts. While Part 1 asks questions about personal and familiar topics, Part 2 requires a structured 2-minute monologue, and Part 3 involves a two-way discussion on abstract issues (British Council et al., n.d.). This final section is widely considered the most challenging part of the test. Test-takers are expected to expand their ideas logically, provide relevant justifications and speculations, and use a range of topic-specific vocabulary and complex grammar structures within a time limit (Le, 2017).

With the development of technology in recent decades, AI, particularly ChatGPT, can offer a promising solution to these challenges by providing simulated conversations, immediate and personalized feedback on fluency, vocabulary, grammar, pronunciation, and idea expansion, and can create a low-pressure practice environment that can reduce general speaking anxiety (Alsalem, 2024; Grab, 2025; Hitokdana, 2025). However, it remains unclear whether this technology can address the cognitive and psychological barriers that Vietnamese learners encounter in dealing with the demanding, highly abstract nature of IELTS Speaking Part 3.

Both existing global and domestic studies have explained ChatGPT's application in IELTS Writing tasks and general speaking practice (Hoang, 2025; Le & Nguyen, 2026; Saricaoglu & Bilki, 2025; Song & Song, 2023; Widyawati, 2026). However, previous studies rarely explore the viewpoints of English-major undergraduates, who are usually trained to analyze linguistic matters critically. Furthermore, no research in Vietnam has been conducted on students' perceptions of AI feedback for high-stakes tests, especially IELTS Speaking Part 3. Therefore, this study aims to bridge the gap by investigating such students' perceptions of ChatGPT-generated feedback for this specific testing context. Understanding learners' experiences, perceived usefulness, limitations, and trust in AI feedback is important for the meaningful integration of such tools into teaching methods. In addition, uncovering these deep personal insights is also likely to help design targeted experimental studies or develop accurate questionnaire items for larger quantitative surveys in the future.

The present qualitative study aims to close this gap by answering the following question:

How do students perceive the value of ChatGPT-generated feedback in supporting different aspects of IELTS Speaking Part 3?

This study is limited to three dimensions. First, it investigates qualitative perceptions, not quantitative score changes. Second, the participants are eight English-major undergraduates from a University in Ho Chi Minh City, who have achieved IELTS band scores of 5.5-6.5. Finally, the study focuses on ChatGPT as the AI tool, excluding others because ChatGPT is highly capable of holding open-ended dialogue and giving contextualized feedback, which are important for simulating an IELTS Speaking Part 3 examiner.

2. Literature Review

2.1. Speaking and its grading criteria

Speaking is defined in different ways. For example, Bygate (1987) describes it as making sounds to interact and get responses from others. Nunan (2003) defines it simply as using spoken words to share meaning. In this study, speaking is understood as a process of making meaning, which involves thinking about ideas and physically expressing them to communicate effectively in a specific situation (Burns & Joyce, 1997).

Speaking is complex and interactive because it involves both linguistic and cognitive factors. Thornbury (2005) states that good speaking requires accuracy in grammar, vocabulary, and pronunciation, fluency, and the use of complex language. Similarly, Common European Framework of Reference for Languages (CEFR) lists vocabulary range, accuracy, fluency, interaction, and coherence as its main parts (Council of Europe, 2020). These criteria align with how the IELTS Speaking test is graded, which focuses on four specific areas, namely Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, and Pronunciation (British Council, n.d.). Therefore, the current study will use these four grading criteria as a guide to explore how learners perceive the AI tool in different aspects of speaking.

2.2. Speaking IELTS part 3

IELTS Speaking Part 3 is a two-way discussion between the examiner and candidate on abstract topics related to Part 2 cue cards. Unlike Part 1 (personal questions on familiar topics) or Part 2 (a 2-minute monologue), Part 3 requires the candidate to think with a higher order, such as idea expansion, argument justification, speculation, comparison, and evaluation (British Council et al., n.d.). The band descriptors clearly favor lexical resource, grammatical range and accuracy, fluency and coherence, and pronunciation (British Council, n.d.).

Vietnamese test-takers often report Part 3 as the most challenging section. They have difficulty in expanding ideas logically, using a wide range of academic vocabulary for unfamiliar and abstract topics, and responding spontaneously without rehearsal (Le, 2017). Rote memorization of model answers, often employed in traditional classrooms, is highly unlikely to facilitate spontaneous, complex discourse, leaving many candidates stuck at band 5.0-6.0 despite strong receptive skills (Vietnamnet, 2024).

2.3. Perception

In language learning, perception is defined as learners' beliefs, attitudes, and evaluations of tools, methods, or learning experiences (Ellis, 2008). According to Dörnyei (2009), learners' perceptions are crucial because they affect participation, persistence, and learning outcomes. In the study by Darma et al. (2023), EFL students' perceptions of technology were explored across three areas: their understanding of the tool, perceived advantages, and perceived disadvantages of

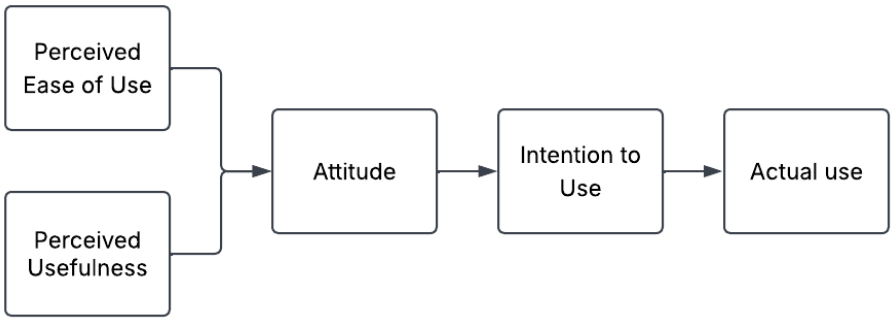
ChatGPT. Therefore, in this study, perception refers to how learners view and evaluate their experiences with ChatGPT-generated feedback for IELTS Speaking Part 3 practice. It includes both the benefits and challenges they perceive while using the tool.

2.4. Technology Acceptance Model (TAM)

This study uses the Technology Acceptance Model (TAM), originally developed by Davis (1989), as the theoretical framework (See in figure 1). The original TAM is made up of five constructs, which are Perceived Ease of Use (PEOU), Perceived Usefulness (PU), Attitude Towards using (AT), Behavioral Intention to use (BI), and Actual Use (AU). Davis (1989) explained that how easy and useful a tool seems (PEOU and PU) shapes a person's attitude toward it (AT). This attitude then influences their intention to use the tool (BI), which ultimately leads them to actually use it (AU).

Figure 1.

TAM from Davis (1989)



While TAM is comprehensive for evaluating technological adoption, this study intentionally narrows down to two main constructs, which are Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). This adaptation is widely accepted by other research in the field. According to Granić and Marangunić (2019), who reviewed TAM in education, researchers often choose PU and PEOU because they are primary independent variables that affect learners' acceptance of a new tool. Similarly, Alsaedi (2025) highlighted that PU is the strongest predictor of the use of AI in learning languages. Due to the time constraints of the current investigation, focusing on PU enables a deeper understanding of ChatGPT's benefits and drawbacks.

This study uses this adapted model for qualitative research. The focus is on PU because it gives insight into how students subjectively weigh its strengths in helping them organize thoughts, develop ideas and vocabulary, fix grammar, and rehearse without pressure during IELTS Speaking Part 3 practice. Using the PU in TAM model allows the evaluation towards the adoption of AI in IELTS practice to be more structured and align with the IELTS four speaking criteria.

2.5. Previous studies

ChatGPT has become more popular in English language learning since its arrival in late 2022. Numerous studies report that ChatGPT is perceived as a useful, accessible, and non-judgemental tool for EFL learners to complete their language tasks. Students appreciate that it can give instant explanations, provide vocabulary, support grammar, and customize exercises (Alghasab, 2025; Alsalem, 2024; Teng, 2024). Many learners also value the less stressful

environment it creates, which encourages them to take risks and develop self-learning (Xiao & Zhi, 2023). However, it raises concerns over abuse, potential inaccuracies, cultural misinterpretation, and academic integrity (Baek et al., 2024; Warschauer et al., 2023). Consequently, to help learners manage these issues while studying independently, Van Horn (2024) and Li et al. (2024) find it important to provide proper training in prompt design and critically evaluate AI responses to maximize benefits while minimizing drawbacks.

Research on ChatGPT for practicing speaking is recent but promising. Studies by Alsalem (2024) and Hitokdana (2025) find that ChatGPT is considered effective in reducing speaking anxiety, increasing motivation, and providing flexibility in practice. The main reason is that AI creates an easy and comfortable environment where learners can practice repeatedly to improve their fluency without worrying about being judged by others. Furthermore, Nguyen and Nguyen (2025) found that using ChatGPT to prepare for IELTS Speaking helped Vietnamese EFL learners develop ideas, expand their vocabulary, and practice independently. As a result, participants reported greater confidence and improved speaking performance.

As for learner engagement, postgraduates using ChatGPT to practice IELTS Speaking increased practice frequency, chose better grammar, and were more engaged, despite having a lack of fully authentic human-like interaction (Zhang et al., 2025). Other studies highlight ChatGPT's ability to provide feedback that is instant and aligns with given criteria, such as fluency, coherence, and vocabulary (Meng, 2025; Sok & Shin, 2025). Limitations often include imperfect speech recognition, generic responses, and reduced trust when feedback feels inauthentic or inaccurate (Alghasab, 2025; Hitokdana, 2025; Jeon et al., 2024).

Previous research shows that ChatGPT helps with general speaking and IELTS writing practice. However, little is known about its use for IELTS Speaking Part 3, especially among Vietnamese English majors. Unlike regular learners, these students are trained to analyze language deeply and discuss complex topics, which makes their views on AI feedback uniquely valuable. To address this gap, this qualitative study explores how they perceive ChatGPT's feedback in this challenging test.

3. Methodology

3.1. Design of the study

To better understand learners' experiences and perceptions, this study adopted a qualitative research design (Creswell & Poth, 2018). According to DeJonckheere and Vaughn (2019), qualitative research is suitable for exploring participants' beliefs, attitudes, and understanding of contexts, including the use of AI tools in language learning. Since the study focused on students' previous experiences, a retrospective qualitative design was used, which allows participants to reflect on their use of ChatGPT.

3.2. Research Context

This study was conducted at an English-Medium Instruction (EMI) university located in Ho Chi Minh City, Vietnam. At this university, students are required to achieve at least an IELTS band score of 5.5 before registering for their major courses. This setting was considered relevant because the participants were already familiar with the IELTS examination and were motivated to improve their speaking ability for academic purposes.

3.3. Participants

Participants were chosen purposely, and they participated voluntarily. The study involved eight English-major undergraduates, including two males and six females, who had IELTS overall band scores ranging from 5.5 to 6.5. Participants needed to have previous experience using ChatGPT to practice IELTS Speaking Part 3 and receive feedback before the research began. After reading the information sheet and providing informed consent, all participants agreed to join voluntarily. Their ages ranged from 21 to 22 years. Creswell and Poth (2018) noted that this type of sample size is suitable because qualitative research focuses on a deep understanding rather than broad results.

3.4. Data Collection Instrument

The study used individual semi-structured interviews to collect data. This interview format is considered flexible because it allows researchers to ask follow-up questions and delve deeper into participants' responses, enriching the data compared with structured interviews (Kallio et al., 2016; Ruslin et al., 2022).

The interview guide mainly focused on the perceived usefulness of the TAM model, examining the benefits and limitations, based on the four IELTS Speaking Part 3 assessment criteria. Since the questions focused on prior experiences, the researchers asked participants to recall their independent practice sessions, describe the prompts they used, provide examples of ChatGPT's feedback, and explain how AI affected their speaking performance and mental readiness.

Individual interviews were used instead of focus groups to allow participants to freely discuss sensitive topics (Gray et al., 2020). All interviews were done through Zoom. This is because it was convenient to schedule and record interviews with participants with their consent while reducing geographical barriers (Gray et al., 2020; Oliffe et al., 2021). Vietnamese was used throughout the interviews, as it was the participants' first language. This helped them express themselves more naturally and reduced misunderstandings that might occur when speaking in a second language (Rolland et al., 2023). Each interview lasted around 30 to 40 minutes.

3.5. Data Analysis

All interviews were audio-recorded with participants' permission and transcribed. The Vietnamese transcripts were first translated into English using ChatGPT. Afterwards, two researchers reviewed and edited the translations independently to check their accuracy and make sure the wording sounded natural.

This study combined Thematic Analysis (Braun & Clarke, 2006) and qualitative Content Analysis (Hsieh & Shannon, 2005). Thematic Analysis was the main method used to identify and organize patterns in the data. The initial codes were deductively created based on two main sources. First, the coding was based on the four IELTS Speaking band descriptors, namely fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation. Second, the codes were derived from themes found in previous studies on language learning supported by AI. As for perceived advantages, codes such as immediate feedback and lexical support were adapted from studies of Meng (2025) and Sok and Shin (2025). The code for grammatical range and accuracy was drawn from existing studies about AI's ability to correct mistakes and give complex

structures at the same time (Kohnke et al., 2023; Su et al., 2023). In a similar vein, idea development and coherence were adapted from studies of Barrot (2023) and Yan (2023), who noted that AI can help generate ideas and expand arguments. The code for a secure practice environment was collected from the studies of Grab (2025) and Hitokdana (2025). In this study, a secure practice environment means having the convenience to practice repeatedly to become fluent. Regarding limitations, the code for speech recognition issues and overreliance on AI were derived from studies of Baek et al. (2024) and Jeon et al. (2024).

However, the analysis was not limited to these codes. New emerging themes were also identified from participants' responses. For example, themes such as prompt quality and learners' critical judgment toward AI feedback appeared during the coding process. Qualitative content analysis was also applied to categorize the findings and count how often they appeared. The mixed coding approach helps the study look at data from both existing frameworks and participants' real experiences (Braun & Clarke, 2006).

3.6. Trustworthiness

The interview questions were reviewed by two experts in TESOL and qualitative research to check whether they were clear and relevant to the study. The researchers also shared summarized findings with two participants so they could confirm that the interpretations reflected their actual experiences.

Interview transcripts, translation notes, and records were maintained throughout the study. In addition, two researchers coded the data independently and discussed any differences until they reached agreement. In addition, detailed information about the research setting, participants, and research procedures was included to help readers consider their application in further research.

4. Results

From the analysis of the semi-structured interviews, two main themes emerged. Those are related to students' perceptions of ChatGPT-generated feedback for IELTS Speaking Part 3. The themes were Advantages Enhancing Perceived Usefulness, Limitations Affecting Perceived Usefulness.

4.1. Advantages Enhancing Perceived Usefulness

Participants strongly emphasized several advantages when discussing how ChatGPT enhances the perceived usefulness of the IELTS speaking part 3. One key finding was the idea of a secure environment to practice. In this study, the majority of participants were satisfied with how accessible ChatGPT was for them to practice more often. P1 explained, "*Since classroom time is limited and valuable, I preferred testing my ideas and answers with AI first*." Similarly, P2 mentioned the good side of repeated practice by saying, *"There are questions I answered 4–5 times and still wasn't satisfied; a real teacher can't wait forever. But with AI, I just practice until I am fluent."* They appreciated being able to practice difficult questions repeatedly at any time without worrying about time limits or judgment from others. P8 shared similar ideas, "*Because nobody laughs and nobody remembers, it helps me dare to try more complex sentence structures that I normally wouldn't dare to use*." This means learners felt more comfortable taking risks when using ChatGPT as they did not worry about being judged. P5 and P7 added similar ideas that unlike human teachers, ChatGPT did not get tired of repeated practice. This means that learners could

practice the same answer many times until they felt more confident and fluent. This also made students become less concerned about wasting class time.

Almost all participants valued the immediate feedback provided by ChatGPT. They felt that receiving feedback right after speaking helped them notice mistakes while their ideas were still fresh in their minds. Some examples include *"because the feedback comes right away, I can still remember what I was trying to say and correct my mistakes more easily"* (P1), *"compared to waiting for teacher feedback, ChatGPT responds much faster, so I can practice more questions in a shorter amount of time."* (P2), and "*what I like most is that ChatGPT gives feedback at once, so I don't have to wait until the next class to know what I did wrong to adjust*" (P5). However, a participant expressed a concern about being too reliant on AI-generated answers, saying, "*I felt like it made me become lazy in coming up with my own ideas."*

For lexical resources, many participants agreed that ChatGPT was useful for supporting paraphrasing. This is because the AI suggested a wide range of alternative vocabulary and collocations for each specific case or for each personalized need (P2, P5, P6). To be more specific, P1 mentioned that he can learn topic-specific vocabulary from ChatGPT, especially in subjects such as technology, education, and the environment, without spending much time searching for this vocabulary online. In addition, P7 and P8 showed a positive attitude related to the fact that ChatGPT helped them overcome the habit of using repeated words by suggesting a variety of words or phrases that can be used in a similar context, with specific examples or even providing more controlled practice in case they felt that they still felt confused about using the words. For example, the participants (P6, P7, and P8) listed out some useful phrases given by ChatGPT to introduce opinions (in my opinion, from my perspective), give reasons (this is mainly because, one reason for this is that), provide examples (for example, for instance), and show contrast (however, on the one side, on the flip side). At the same time, they received sample sentences on how to use these phrases from ChatGPT. Benefits related to collocations were also stated by P1 and P4. They reported that before using ChatGPT, their knowledge of collocations was limited because materials on collocations were scarce.

ChatGPT was also reported to support participants with complex sentence structures. P1, P2, and P4, respectively, said that "*ChatGPT tends to replace my simple sentences with complex sentences*", "*it often gives me structures like although, while it is true that*", and "*conditional sentences are one of the most common structures I get from ChatGPT*". While P5 liked the idea that ChatGPT could help suggest varied structures in speaking Part 3, especially for opinion questions, P6 emphasized the awareness that ChatGPT raised for the practitioner when giving feedback to the grammatical resources and the suggested versions, which include different grammatical structures. P7 also noted that ChatGPT could help buy time before answering challenging prompts by using "*That's an interesting question*".

In terms of idea development and coherence, most participants appreciated ChatGPT for helping them develop ideas logically and connect their ideas cohesively. P3 said that "*the AI helps me develop the supporting details for my arguments by suggesting ideas related to explanations, causes or effects, which directs me to expand relevant ideas faster*". P4 reported that "*ChatGPT helps me think about the good and bad sides of things with specific explanation or examples or further consequences*", while P6 noted "*ChatGPT opens my mind since it considers new viewpoints which often followed by explanation*". P7 liked the way ChatGPT suggested to arrange

the two main ideas of each response in Speaking Part 3 from less important to more important or from less serious to more serious, which gradually shaped the logic of her speech. In addition, to ensure cohesion, the AI often provided linking words and discourse markers. P1 and P2 shared similar experiences, saying that "*I often receive words like however, for example, on the other hand, and on top of that, etc. to add to my answers*". In most of ChatGPT's feedback, cohesive devices were mentioned explicitly, reminding the speaker to be aware of linking ideas effectively. For example, P3 shared that to strengthen the ideas or organize the ideas in the response more logically, phrases such as one of the reasons is that, another reason is that, one possible solution can be, government should, etc. were often suggested.

4.2. Limitations Affecting Perceived Usefulness

Participants faced limitations that affected ChatGPT's usefulness, especially with prompt quality and speech recognition. P2, P6, and P7 reported that the quality of feedback depended a lot on how the prompts were written. P3 estimated that prompt quality could account for from 80 to 90% of the feedback's usefulness, noting, "*The prompt is basically what you give ChatGPT so it knows what you want it to do*." P8 also emphasized that "*a good prompt is what differentiates me from others using the same AI.*" When prompts were too broad or unclear, the feedback often felt general and not specific enough to be helpful. Because of this, participants gradually created their own ways of using the tool more effectively. For example, P2 learned to ask ChatGPT to evaluate responses based on IELTS band descriptors, while P6 spent time trying different prompt styles to see which ones produced better feedback.

Participants also shared different experiences with speech recognition. P3, P6, and P8 did not report any serious problems. Meanwhile, P1 and P2 were frustrated with occasional technical lag and speech misrecognition. This is because they interrupted their speaking flow and made it harder to maintain their ideas. However, instead of stopping, P5 viewed it as a "*challenge to adjust and pronounce more clearly and accurately*" Interestingly, P4 preferred typing instead of using the voice feature to avoid these problems completely.

In addition, participants realized that using ChatGPT could be a double-edged sword. While it provided feedback instantly, they recognized that they might become too dependent on it. P1 said, "*if I rely too much on ChatGPT, my brain just stops working*" when she found herself waiting for ChatGPT's answers instead of her own ideas first. P2 and P4 also shared similar ideas, such as "*I might lose my ability to think without the help of ChatGPT*" and "*I sometimes feel really lost when not using AI*". Because of this overreliance, learners tend to lack independent thinking. Consequently, they might blindly adopt all AI's suggestions, even when those cues are not accurate. Therefore, participants knew they needed to think critically about ChatGPT's feedback. Many did not fully trust the AI because they had received wrong or unreliable information. For instance, P3 and P6 reported encountering situations in which ChatGPT provided incorrect information or unverifiable links. P6 clearly said "*I've seen ChatGPT create fake links and wrong information."* Similarly, P8 mentioned that the AI sometimes suggested collocations that sounded outdated or changed the original meaning when correcting grammar.

As a result, participants had a tendency to double-check the feedback instead of accepting it immediately. P2 relied on reliable dictionaries such as Cambridge and Oxford, while P5, P6, and P7 checked AI-generated information against research articles and other online sources. These experiences suggest that participants generally saw ChatGPT as a tool to support learning rather

than something that could fully replace human teachers. P6 mentioned, " *My trust in ChatGPT is never 100%; I always need to check again."*

5. Discussion

5.1. Advantages Enhancing Perceived Usefulness

In line with the Perceived Usefulness construct of the TAM model, the qualitative data showed that ChatGPT helps reduce the cognitive and psychological load of IELTS Speaking Part 3. The majority of the participants (5 out of 8) found ChatGPT useful, especially because it created a secure environment for practice and provided immediate feedback. When using this tool, they are not afraid of being judged by a real person, which prevents them from losing face or feeling hurt. Instead, they can overcome anxiety when speaking and produce more coherent ideas or more fluent speech, thanks to the convenience of repeating answers. This preference is largely attributed to the demanding, spontaneous nature of Part 3, which often causes speaking anxiety. Similar findings were reported by Grab (2025) and Hitokdana (2025), who confirmed that removing human judgment can eliminate the fear of making mistakes and encourage students to keep practicing. In addition, participants particularly liked the fact that ChatGPT is tireless. For the same number of students, this flexibility was important since classroom practice is often limited by time and class size. This comfort can motivate them to practice frequently, reconsider their responses, recognize their problems, and adjust and revise them. They could also attempt more complex sentence structures until they felt more confident and fluent, or until they reached the day's practice time limit. Therefore, when using ChatGPT, learners can actually improve their speaking performance and confidence. This finding is consistent with Zhang et al. (2025), who reported that learners using ChatGPT practiced more frequently and became more engaged in IELTS preparation.

The findings also showed that most of the participants (6 out of 8) found ChatGPT helpful in suggesting a wider vocabulary range and in paraphrasing to avoid repetition when responding to the Speaking Part 3 question. In addition, collocations were a plus point reported by those users, helping them achieve a higher band score. This effectiveness may stem from ChatGPT's design as a generative AI model. Unlike traditional language-learning apps that offer only pre-programmed phrases, ChatGPT can analyze learners' natural, spontaneous speech and immediately suggest suitable alternatives based on the provided prompts, with detailed descriptions of the criteria and examples (Kohnke et al., 2023). Receiving these corrections while their original thoughts remain active in memory eases the process of internalizing advanced academic words or paraphrasing to avoid repetitive language use. As a result, the revision process was faster. Furthermore, ChatGPT also benefited participants by offering useful phrases to introduce opinions, give reasons and examples, show contrast, etc. (3 out of 8). Since AI also provided sample sentences and controlled practice, learners were able to use new expressions effectively (2 out of 8). These findings align with the results from Meng's (2025) and Sok and Shin's (2025) studies, which found that ChatGPT was useful for expanding topic-specific vocabulary and collocations, as well as helping learners avoid repeating words. This real-time, personalized language support helps students articulate abstract ideas in Part 3 with more confidence.

Furthermore, the interview findings highlighted the roles of ChatGPT in assessing and expanding learners' grammatical range. In the beginning, it raised participants' awareness of grammatical variety by evaluating their spoken responses against the IELTS criteria. For example,

P6 reported that this detailed feedback helped them realize their structural limitations. Three participants (P1, P2, P4) valued the AI's ability to diversify their sentence structures by replacing their simple sentences with more complex, content-specific alternatives such as conditional tenses or concession clauses. As highlighted by Teng (2024), this AI feature works well in this regard because it personalizes support to fit the context. Moreover, studies by Kohnke et al. (2023) and Su et al. (2023) found that AI helps correct mistakes and develop structures simultaneously based on the learner's original ideas. Although these studies focused on writing, the positive findings regarding correcting and expanding grammatical structures are similar. As the diversity of grammatical resources is one of the criteria for improving IELTS speaking band scores, this benefit is worth considering.

Finally, the findings showed that 6 out of 8 participants found that ChatGPT was effective at providing comments on content development. They also reported that ChatGPT helped them a lot improve the overall structure and the logical connections in their spoken responses. While Meng (2025) and Sok and Shin (2025) noted that ChatGPT provides helpful feedback on coherence, this study builds on this by showing how the AI does so in IELTS Speaking Part 3 practice. Specifically, to build this structural coherence, ChatGPT encouraged participants to develop their answers by explaining causes and effects, looking at advantages and disadvantages, and considering various viewpoints when answering Part 3 questions. This finding also supports Barrot (2023), who argued that ChatGPT helps EFL learners generate and expand their ideas. Even though this study focuses on the evaluation stage rather than the pre-stage, ChatGPT also plays a role in suggesting additional ideas to help learners think critically and select the most suitable one. Similarly, Yan's (2023) findings confirmed that ChatGPT assists learners in broadening their arguments, explaining them in greater detail, and exploring multiple perspectives. Although Barrot (2023) and Yan (2023) focused on second-language writing, this study shows that their findings also apply to IELTS Speaking Part 3 practice, where the learner's competence is expressed by the ability to provide strong arguments with relevant supporting details to convince the examiner. Although AI cannot replace human teachers, it can be a useful tool for IELTS Speaking Part 3 practice.

5.2. Limitations affecting Perceived Usefulness

Despite these benefits, the study shows that ChatGPT's usefulness was affected by technical limits and prompt design, reflecting the Perceived Ease of Use in the TAM framework. Most participants (5 out of 8) identified a primary limitation, which was prompt dependency. The gap in feedback quality stems from the fact that AI does not naturally know how to teach. If students do not give specific instructions, such as asking the AI to use IELTS band descriptors and to provide specific examples for well-qualified responses, along with guidance on how to comment on the response, ChatGPT simply gives vague, generic feedback, as noted by those five participants. The tool becomes less user-friendly when students must put a lot of mental effort into figuring out the perfect prompts. As Van Horn (2024) and Li et al. (2024) argued, students need training in prompt design to achieve the best possible results, as vague feedback often erodes learners' trust and interest. Therefore, to learn independently with ChatGPT, students must understand how to communicate effectively with AI by equipping themselves with AI principles and effective prompts.

Another problem was speech misrecognition, reflecting the technical barriers mentioned by Jeon et al. (2024). Half of the participants in this study felt frustrated when the app lagged or misheard their accents, mainly because this interrupted their train of thought. However, instead of giving up, learners actively adapted to these technical flaws. One student (P4) skipped the voice feature and switched to typing to receive more accurate feedback on the quality of their responses, excluding pronunciation, voicing techniques, and fluency, while the other (P5) treated the flaws as a chance to practice speaking more clearly. This shows that motivated learners will find ways to make the tool work for them despite speech misrecognition. From these efforts, they could improve their speech performance, especially by adjusting their pronunciation and voice clarity.

Besides that, several students (P1,P2,P4) were worried about becoming too dependent on ChatGPT because it sometimes provided incorrect information or unnatural phrases. Because ChatGPT sometimes makes up information or uses words out of context, it raises concerns about accuracy and academic integrity, as pointed out by Baek et al. (2024). Consequently, if students accept these responses without thinking critically, they might risk losing their critical thinking skills in the long run (Warschauer et al., 2023). To avoid this, nearly all students (6 out of 8) in this study adopted a strategy of always double-checking the AI's comments using their knowledge of grading criteria, reflective and critical thinking skills, and the foundations of language components or discourses. For comments that seemed confusing, P5, P6, and P7 shared that they would use AI to suggest reliable, relevant sources to read about the issue they are facing, so that they can judge if AI provides accurate comments. This supports Van Horn (2024) and Li et al. (2024), who argued that critically evaluating AI responses is absolutely necessary, emphasizing that learners need to actively judge and use their reflective skills to filter, verify, and choose the most accurate AI’s responses. Although checking facts takes time, it actually helps students become more independent, noted by Xiao and Zhi (2023).

Furthermore, this situation redefines the teacher's role. As Teng (2024) described, ChatGPT is a "companion, not enemies." Since AI can handle repetitive practice and basic vocabulary, human teachers are free to focus on higher-level thinking. Human instructors cannot be replaced because AI cannot replicate real human conversation or understand context perfectly. Teachers are essential not only for verifying AI feedback, but also for providing the real conversational pressure and emotional support needed for IELTS Speaking Part 3. Ultimately, ChatGPT is a powerful support tool, but it works best when combined with a human teacher's expertise.

6. Conclusion

This study explores undergraduates’ perceptions towards the use of ChatGPT’s feedback for IELTS Speaking Part 3 practice. Overall, the participants found ChatGPT to be highly convenient and useful for test practice. Without being rushed or judged, they appreciated ChatGPT for creating a secure environment where they could practice frequently and repeatedly until they felt truly fluent. The tool also helped them improve their speaking by giving instant feedback and suggesting advanced vocabulary.

However, students also found some difficulties when using the tool. Many struggled to write quality prompts, which often led to unhelpful feedback from the AI. Speech recognition errors were also quite common, which disrupted their speaking flow and caused frustration. In addition, students had difficulty in applying the advanced vocabulary suggested by ChatGPT naturally into

their speech. This is usually because the AI could not explain the specific context of the words like a real human instructor would.

These findings offer practical suggestions for both teaching and learning. Using ChatGPT for practicing fluency with lots of repetition and vocabulary building can be very effective. However, it needs to be combined with human guidance. Teachers should train students on how to write specific prompts and how to critically evaluate the AI's feedback, so they do not simply accept incorrect information. Because ChatGPT cannot act exactly like a real person in a conversation, human teachers are still essential. They are needed to verify the AI's information and guide students to improve their reflexes needed in the real IELTS Speaking test.

This study, however, remains limited. This is because the research only involved a small group of students from one university. Consequently, the results might not represent all EFL learners. Moreover, the study only focused on students' opinions and feelings, not their actual test scores. For future research, it would be helpful to study whether using ChatGPT actually improves students' IELTS Speaking scores over a longer period. It will also be valuable to train students to write better prompts because these can affect their learning outcomes.

Acknowledgements

We would like to express our sincere gratitude to all the students who participated in this study. Your honest responses and valuable time greatly contributed to the success of this research. Without your cooperation and generosity, this research would not have been possible. I would also like to thank the conference chairs and reviewers for their helpful comments on earlier drafts of the article.

References

Alghasab, M. (2025). *Insights into EFL students' perceptions of the "ChatGPT essentials" training course for language learning*. *Education Sciences, 15*(9), 1138. https://doi.org/10.3390/educsci15091138

Alsaedi, N. S. (2025). Exploring ChatGPT's role in EFL learning through the technology acceptance model: Perspectives from Saudi students. *Contemporary Educational Technology*, *17*(4), ep594. https://doi.org/10.30935/cedtech/17302

Alsalem, M. S. (2024). EFL students' perception and attitude towards the use of ChatGPT to promote English speaking skills in the Saudi context. *Arab World English Journal, 15*(4), 73–84. https://doi.org/10.24093/awej/vol15no4.5

Baek, C., Tate, T., & Warschauer, M. (2024). "ChatGPT seems too good to be true": College students' use and perceptions of generative AI. *Computers & Education: Artificial Intelligence, 7*, 100294. https://doi.org/10.1016/j.caeai.2024.100294

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

British Council. (n.d.). *Speaking band descriptors*. https://takeielts.britishcouncil.org/sites/default/files/ielts_speaking_band_descriptors.pdf

British Council, IDP IELTS, & Cambridge Assessment English. (n.d.). *IELTS test format*. https://www.ielts.org/

Burns, A., & Joyce, H. (1997). *Focus on speaking*. National Centre for English Language Teaching and Research.

Bygate, M. (1987). *Speaking*. Oxford University Press.

Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *Tesol quarterly, 48*(4), 738-762. https://doi.org/10.1002/tesq.148

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe Publishing.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Darma, R. N., Syahid, A., Fatma, Lestari, A. P., Basahil, A., & Ningrum, R. L. (2023). The student's perception of using ChatGPT for EFL students. *Jurnal Ilmu Pendidikan Nasional (JIPNAS), 1*(3), 143–146. https://doi.org/10.59435/jipnas.v1i3.193

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3)*,* 319-340. https://doi.org/10.2307/249008

DeJonckheere, M., & Vaughn, L. M. (2019). Semistructured interviewing in primary care research: A balance of relationship and rigor. *Family Medicine and Community Health, 7*(2), e000057. https://doi.org/10.1136/fmch-2018-000057

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Duong, T. H. O., & Nguyen, T. H. (2006). Memorization and EFL students' strategies at university level in Vietnam. *TESL-EJ, 10*(2). http://files.eric.ed.gov/fulltext/EJ1065021.pdf

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.

Grab, M. Ö. (2025). Integrated AI chatbot practice: A pathway to improved ESL speaking skills. *Social Sciences & Humanities Open, 12*, 101933. https://doi.org/10.1016/j.ssaho.2025.101933

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. British Journal of Educational Technology, 50(5), 2572-2593. https://doi.org/10.1111/bjet.12864

Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *The Qualitative Report, 25*(5), 1292-1301. https://doi.org/10.46743/2160-3715/2020.4212

Hitokdana, K. O. (2025). Students' perceptions of the benefits and challenges of using ChatGPT for English speaking: A systematic review. *ELS Journal on Interdisciplinary Studies in Humanities, 8*(4), 1209–1221. https://journal.unhas.ac.id/index.php/jish/article/view/48815

Hoang, N. H. (2025). ChatGPT-assisted language learning: Effects on Vietnamese English majors' writing skills and motivation. *The JALT CALL Journal, 21*(2), 102429. https://doi.org/10.29140/jaltcall.v21n2.102429

Van Horn, K. R. (2024). ChatGPT in English language learning: Exploring perceptions and promoting autonomy in a university EFL context. *TESL-EJ, 28*(1). https://doi.org/10.55593/ej.28109a8

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Jeon, J., Lee, S., & Coronel-Molina, S. M. (2024). Rethinking AI: Bias in speech-recognition chatbots for ELT. ELT journal, 78(4), 435-445. https://doi.org/10.1093/elt/ccae035

Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing, 72*(12), 2954-2965. https://doi.org/10.1111/jan.13031

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Le, P. T. A. (2017). Challenges faced by Vietnamese learners in IELTS speaking: More than just language problems. *Humanising Language Teaching, 19*(5). https://old.hltmag.co.uk/oct17/mart03.htm

Le, T. N., & Nguyen, H. H. (2026). Vietnamese EFL teachers' perspectives on using ChatGPT to support IELTS writing instruction. *Issues in Educational Research, 36*(1), 137-160. http://www.iier.org.au/iier36/le-abs.html

Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers & Education: Artificial Intelligence, 7*, 100266. https://doi.org/10.1016/j.caeai.2024.100266

Meng, F. (2025). Customizing ChatGPT for second language speaking practice: genuine support or just a marketing gimmick? *Proceedings of the International Conference of the Learning Sciences*, 2056–2060. https://doi.org/10.22318/icls2025.615428

Muluk, S., Zainuddin, Z., & Dahliana, S. (2022). Flipping an IELTS writing course: Investigating its impacts on students' performance and their attitudes. *Studies in English Language and Education, 9*(2), 591-612. https://doi.org/10.24815/siele.v9i2.23314

Nga, T. (2024, October 14). Vietnam's average IELTS score falls in 2023–2024, only 1% score above 8.5. *VietNamNet*. https://vietnamnet.vn/en/vietnam-s-average-ielts-score-falls-in-2023-2024-only-1-score-above-8-5-2331801.html

Nguyen, H. N., & Nguyen, D. K. (2022). Vietnamese learners' performance in the IELTS writing task 2: Problems, causes, and suggestions. *International Journal of TESOL & Education, 2*(1), 170-189. https://doi.org/10.54855/ijte.222111

Nguyen, P., & Nguyen, T. (2025). Using ChatGPT to develop students' IELTS Speaking Skill in an IELTS Course. *Proceedings of VietTESOL International Convention 2024: ELT in the Age of Artificial Intelligence, 5*, 92-119. https://proceedings.viettesol.org.vn/index.php/vic/article/view/256

Nunan, D. (2003). *Practical English language teaching*. McGraw-Hill.

Oliffe, J. L., Kelly, M. T., Gonzalez Montaner, G., & Yu Ko, W. F. (2021). Zoom interviews: Benefits and concessions. *International Journal of Qualitative Methods, 20*, 16094069211053522. https://doi.org/10.1177/16094069211053522

Rolland, L., King, H. M., & Lorette, P. (2023). Methodological implications of participant and researcher multilingualism: making language dynamics visible. *Journal of Multilingual and Multicultural Development, 44*(8), 645-656. https://doi.org/10.1080/01434632.2023.2224774

Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F., & Syam, H. (2022). Semi-structured interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education, 12*(1), 22–29. https://doi.org/10.9790/7388-1201052229

Saricaoglu, A., & Bilki, Z. (2025). The capacity of ChatGPT-4 for L2 writing assessment: A closer look at accuracy, specificity, and relevance. *Annual Review of Applied Linguistics*, *45*, 253-273. https://doi.org/10.1017/S0267190525100160

Sok, S., & Shin, H. W. (2025). Do interactions with ChatGPT influence L2 learners' oral speaking ability, summarization ability, and perceptions of generative AI tasks? *TESOL Quarterly, 59*(S1), S19-S51. https://doi.org/10.1002/tesq.70001.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, Article 100752. https://doi.org/10.1016/j.asw.2023.100752

Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers & Education: Artificial Intelligence, 7*, Article 100270. https://doi.org/10.1016/j.caeai.2024.100270

Thornbury, S. (2005). *How to teach speaking*. Pearson Education.

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing, 62*, 101071. https://doi.org/10.1016/j.jslw.2023.101071

Widyawati, A. (2026). A systematic review of teachers' perspectives on using Chat GPT to teach speaking: Benefits and challenges. *PROJECT (Professional Journal of English Education), 9*(1), 261-272. https://journal.ikipsiliwangi.ac.id/project/article/view/30014

Wu, W.-C. V., Yang, J. C., Hsieh, J. S. C., & Yamamoto, T. (2019). Free from demotivation in EFL writing: The use of online flipped writing instruction. *Computer Assisted Language Learning, 33*(4), 353-387. https://doi.org/10.1080/09588221.2019.1567556

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3), 212. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zhang, H., Su, Y., Xu, X., Lei, V. N. L., & Hao, S. (2025). Beyond theory: A mixed-methods exploration of postgraduate engagement with ChatGPT for English speaking proficiency. *International Journal of Technology in Education, 8*(4), 1022-1047. https://doi.org/10.46328/ijte.1221

Bionote

Diem Bich Huyen Bui is currently a full-time lecturer and associate of student affairs at the School of Languages, International University - Vietnam National University, HCMC. She completed her undergraduate studies in TESOL and earned her MA in Applied Linguistics from La Trobe University, Australia. She is currently pursuing her PhD in TESOL. Her research interests include formative assessment, learner autonomy, skills development, teaching methodologies, and technology-enhanced language learning (TELL).

Email: bdbhuyen@hcmiu.edu.vn

Phone number: 0918 504 964

Vo Man Nghi Nguyen is currently an undergraduate majoring in English Linguistics and Literature at International University - Vietnam National University, HCMC. She is an IELTS tutor who works with students from different proficiency levels. She is interested in language teaching and research. Her research interests include IELTS preparation, language assessment, technology-enhanced language learning (TELL), and artificial intelligence in English language education.

Email: patricianghi23112004@gmail.com

Phone number: 0977 567 314

#120 DIGITAL TRANSFORMATION IN FORMATIVE ASSESSMENT: THE DESIGN, IMPLEMENTATION, AND EVALUATION OF TESTMOZ FOR ENGLISH GENERAL COURSES AT VIETNAM AVIATION ACADEMY

Trinh Phan Tu[1]

Abstract

The ongoing digital transformation in higher education has exposed limitations in traditional self-study practices at Vietnam Aviation Academy (VAA), particularly as self-directed learning accounts for up to 30% of course assessment. In General English courses, paper-based homework frequently results in delayed feedback, increased administrative workload for instructors, and low levels of student engagement. Despite national efforts to promote Industry 4.0 integration in Vietnamese higher education, homework and self-study activities at many universities in general and VAA in particular remain largely manual and insufficiently aligned with digital learning environments.

To address challenges, the author needs streamlined digital tools that support rather than complicate the learning process. This study investigates the integration of Testmoz as a solution to modernize formative assessment. Specifically, it examines students' perceptions of Testmoz's contribution to feedback quality, learning motivation, homework completion, learner engagement, and autonomous learning. A mixed-methods design following the ADDIE model was employed. Data were collected from 182 undergraduate students who had used Testmoz for formative homework activities over one academic semester through a questionnaire comprising Likert-scale items and open-ended questions. Quantitative data were analyzed descriptively, while qualitative responses provided additional insights into students' learning experiences.

Moreover, the research aim is to clarify if digitizing these practices from designing Testmoz assignment can be able to not only provide students with more timely feedback but also reduce the daily workload for instructors. By implementing and evaluating Testmoz at the Vietnam Aviation Academy (VAA), the research also hopes to provide a practical roadmap for using digital tools to improve student outcomes in the Vietnamese higher education context.

***Keywords*:** digital transformation, Testmoz, assessment, feedback, student perceptions

1. Introduction

In the age of Industry 4.0 development, the digital transformation of many society fields including education has become a national priority in Vietnam. However, a profound gap remains

[1] *Foreign Language Faculty, Vietnam Aviation Academy, Ho Chi Minh City, Vietnam; Email: trinhpt@vaa.edu.vn*

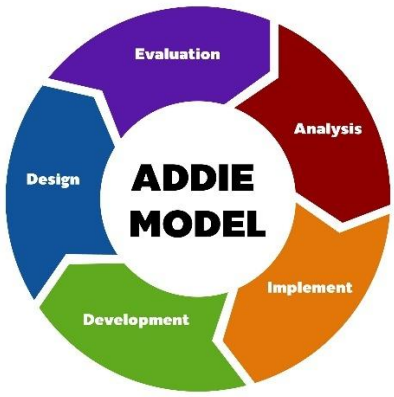
between traditional assignment and classroom reality. At the Vietnam Aviation Academy (VAA) in addition, self-directed learning contributes 30% of the total course assessment, yet this significant component is still largely managed through traditional, paper-based methods. In General English courses, these manual processes often result in "feedback delays," where the gap between a student's effort and a teacher's response is too wide to be effective. This not only increases the administrative burden on instructors but also distracts student motivation.

According to the previous research on digital assessment of management systems (LMSs), It has primarily focused on comprehensive , online learning environments, or large-scale digital assessment platforms (Priyeti et al., 2025). While these studies demonstrate the potential of technology to support teaching and learning, less attention has been paid to lightweight, low-cost assessment tools that can be easily implemented in resource-constrained educational contexts. In particular, few studies have documented how such tools can be systematically designed, implemented, and evaluated using an instructional design framework within General English courses (Pariska & Rayusda, 2026).

To bridge this gap, an urgent need for streamlined digital tools is necessary that they can enhance, rather than complicate, the formative assessment process. This study has concentrated on investigating the integration of Testmoz, a digital platform which is lightweight and versatile, as a strategic solution with the combination of the ADDIE model from Culatta (2018) includings Analysis, Design, Development, Implementation, and Evaluation, the research documents the transition from paper-based homework to a digitized workflow.

Figure 1

ADDIE model



By digitizing homework and self-study activities, the research explores whether Testmoz can facilitate immediate feedback, alleviate the heavy workload of manual grading, and better align the curriculum with digital learning environments. Through the design, implementation, and evaluation of Testmoz at VAA, this study aims to provide a practical roadmap for educators seeking to modernize assessment practices and improve student outcomes by utilizing the digital transformation. From the certain purposes that are to measure the technical success and the "speed" of the new system, to identify the "human" benefit for teachers and to understand the psychological and professional impact on the learners themselves, the paper focuses on addressing 2 research questions.

Research Questions

1. How can Testmoz be designed to align with the General English syllabus at VAA, and to what extent does this implementation improve the timeliness of feedback and student completion rates?

2. How do students perceive the transition from manual to digital assessment, particularly regarding their motivation and control over their own learning?

2. Literature Review

This literature review has researched on digital transformation and formative assessment in higher education, with emphasis on feedback, instructional efficiency, and learner self-study. It also discusses digital assessment tools and instructional design frameworks, identifying a research gap in the ADDIE-based implementation of friendly platforms such as Testmoz in Vietnamese General English courses.

2.1 Digital Transformation in Higher Education

Digital transformation has been widely being recognized as a critical method of innovation and quality improvement in higher education systems worldwide. According to Selwyn (2016), digital transformation in education is not merely about adopting technology but about reshaping pedagogical practices, assessment models, and institutional cultures. In the current era of Industry 4.0, universities are increasingly required to integrate digital tools to enhance flexibility, student autonomy, and efficiency (Dang et al., 2024).

In Vietnam, the Ministry of Education and Training (MOET) has prioritized digital transformation as a strategic impression, particularly after the COVID-19 pandemic which is accelerated the adoption of online and blended learning models (Nhung et al., 2025). However, several researches show that while digital platforms are often used for content delivery, assessment practices, especially formative assessment, remain predominantly traditional and paper-based (Phuong et al., 2024). This mismatch reduces the overall effectiveness of digital learning environments limits students' access to timely feedback and opportunities for self-regulated studying.

2.2 Formative Assessment and the Role of Feedback

The formative assessment plays extremely significantly in supporting learning by giving feedback provision that helps learners identify gaps between current and desired performance (Black & Wiliam, 1998). The quick and effective formative feedback is featured by timeliness, clarity, and actionability (Hattie & Timperley, 2007). When delaying in feedback, as is common in paper-based homework systems, its pedagogical value is crucialy diminished.

Digital formative assessment tools have been clarified to address this issue by permitting immediate or rapid feedback, thereby enhancing student engagement and learning outcomes (Nicol & Macfarlane-Dick, 2006). In language learning contexts, online quizzes and automated

assessment systems are particularly effective in reinforcing grammar, vocabulary, and reading comprehension skills (Chapelle & Douglas, 2006).

2.3 Digital Assessment Tools and Instructor Workload

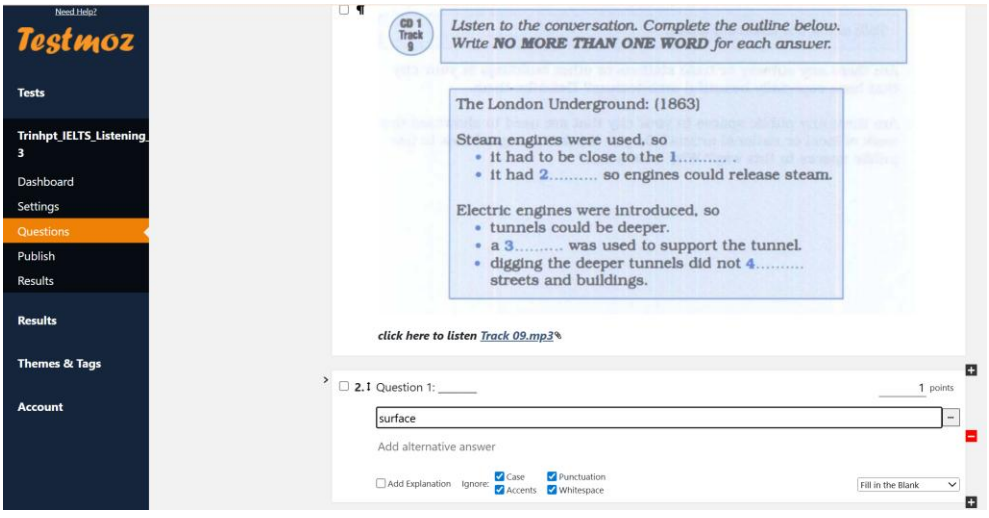
The increasing administrative workload associated with assessment and grading is one of the major challenges that instructors in higher education face nowadays. The study of "Changing assessment towards a new assessment paradigm using ICT" indicates that automated and semi-automated digital assessment tools can significantly reduce grading time, minimize repetitive tasks, and allow instructors to focus more on pedagogical decision-making (Redecker & Johannessen, 2013).

Platforms such as Moodle quizzes, Google Forms, Kahoot, and Testmoz have been widely adopted due to their ease of use and automatic scoring features (Bennett, Dawson, Bearman, Molloy, & Boud, 2017). Lightweight tools like Testmoz are particularly suitable for educators with the limitation of the technical infrastructure, as they do not require complex learning management systems while still supporting large-scale formative assessment.

For specialized assessments such as rearrangement or sentence transformation within the "Fill in the Blank" format, Testmoz offers rigorous evaluation settings. To ensure exact grading and linguistic accuracy, teachers can set four distinct validation modes: 1. Case Sensitivity: Distinguish between upper and lowercase characters; 2. Accents: Validate the correct use of diacritics; 3. Punctuation: Require exact matches for all grammatical symbols; 4. Whitespace: Monitor the specific spacing between words or characters. It has been used in diverse educational settings due to its ease of use and versatility.

Figure 2

Testmoz interface displaying the four validation settings



Unlike some other quiz platforms, Testmoz does not require students to create accounts just click on the link type their name and can easily get access to do the homework, which reduces barriers to participation and is very friendly to use. Having used Testmoz for over 5 years since

the epidemic, the author has found that Testmoz supports continuous assessment and improves homework submission rates. Moreover, its ability to monitor student responses and generate performance reports helps educators to identify areas where learners may need some more additional support.

Figure 3

Student login interface on Testmoz highlighting account-free access via direct link and name entry

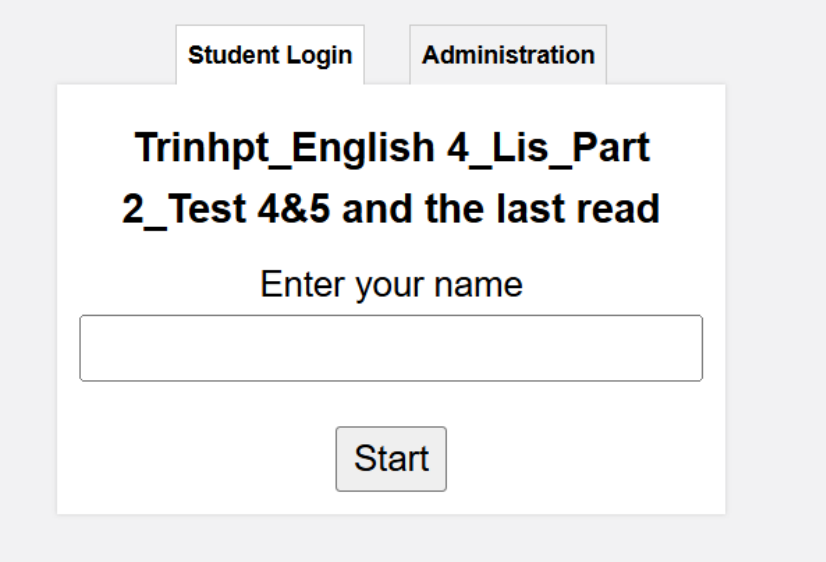
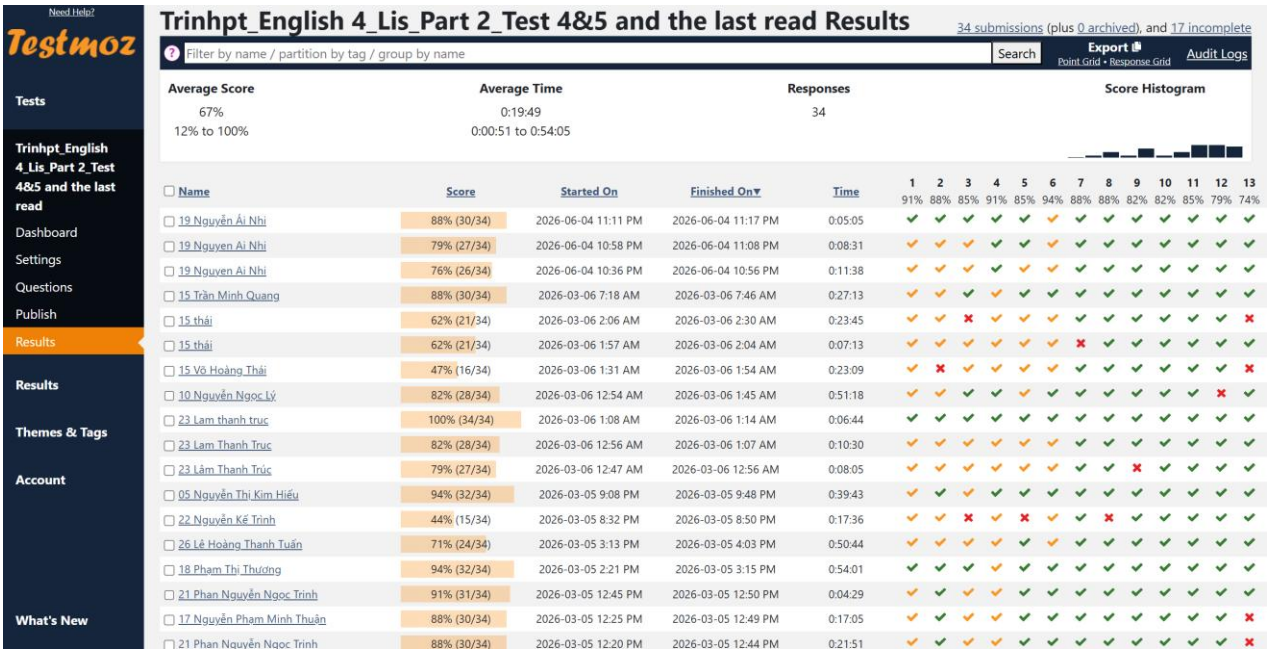


Figure 4

Testmoz administrative dashboard displaying aggregated scores, completion times, and individual item responses for educator analysis



2.4 Student Motivation, Autonomy, and Perceptions of Digital Assessment

Learner perception is a key factor which influences the success of digital assessment implementation. Studies grounded in Self-Determination Theory suggest that digital tools are able to enhance motivation by supporting autonomy, competence, and control over learning pace (Deci & Ryan, 2000). Having immediate feedback and being able to monitor the assignment progress, students can be more likely to take responsibility and interest for their learning (Zimmerman, 2002).

In the Vietnamese EFL context, Nguyen and Dang (2022) have found that students respond positively to online quizzes and digital homework, citing increased convenience, reduced anxiety, and greater transparency in assessment. However, successful adoption depends on clear alignment with course objectives and adequate guidance from instructors.

2.5 Research Gap

While existing literature highlights the benefits of digital formative assessment, there is limited empirical research focusing on lightweight digital tools such as Testmoz in Vietnamese higher education, particularly in General English courses where self-study accounts for a substantial proportion of course assessment. Moreover, few studies adopt a systematic instructional design framework such as the ADDIE model to document the design, implementation, and evaluation of such tools. This study seeks to address this gap by examining

both technical efficiency and human-centered impacts on instructors and students at Vietnam Aviation Academy.

3. Methodology

3.1 Research Design

The study combined the ADDIE instructional design model with a convergent mixed-methods approach, in which quantitative and qualitative data were collected at the end of the implementation and interpreted together to evaluate students' perceptions of Testmoz. The project followed the ADDIE model:

- Analysis
- Design
- Development
- Implementation
- Evaluation

The study involved 182 undergraduate students enrolled in General English courses at Vietnam Aviation Academy. Participants represented multiple English classes and completed the survey voluntarily after using Testmoz for formative homework activities over one academic semester. Ramdomly, the proportion of female students and male students accounted for 50% with 91 students for each gender. Smartphones emerged as the most frequently used device for accessing Testmoz with the percentage of 53%, followed by laptops with 45%, indicating the mobile-friendly nature of the platform.

The questionnaire consisted of Likert-scale items measuring ease of use, engagement, motivation, feedback effectiveness, homework consistency, and continuance intention. Open-ended questions explored learner experiences and suggestions. Internal consistency reliability could be assessed using Cronbach's Alpha to ensure acceptable reliability of the survey instrument.

3.2 Context and Participants

Participants include 182 undergraduate students enrolled in General English courses at VAA. These students, taking English 3 or 4, are familiar with Testmoz and have used it for at least one semester on laptops, smartphones, or tablets to complete quizzes as homework with not only reading and listening but also writing and speaking assignment.

3.3 AI-Supported Assessment Model

Digital quizzes were developed on Testmoz. Features included:

- Automated scoring
- Immediate feedback display
- Randomized question order
- Time limits
- Exportable performance reports

While not generative AI, the system automated assessment processes traditionally performed manually. AI writing tools were discussed in class within ethical guidelines, emphasizing transparency and proper use.

3.4 Ethical Considerations

Participation was voluntary. No personally identifiable information was published. Data were anonymized. Students were informed about data usage and consented prior to survey participation.

4. Findings

4.1 Student Perceptions of Testmoz

Based on the survey results, the data was exported to Excel and categorized into three groups: scores of 1 and 2 were classified as 'Disagree', 3 as 'Neutral', and 4 and 5 as 'Agree'. This grouped data was then analyzed to produce the results shown in Table 1 below:

Table 1

Student Perceptions of Testmoz (N = 182)

Dimension	Disagree (%)	Neutral (%)	Agree (%)	Interpretation
Ease of Use	4.5	8.4	87.1	The vast majority of students found Testmoz easy to use, with only a small proportion reporting difficulties. This indicates that the platform was highly accessible and user-friendly for most learners.
Engagement	5.6	9.6	84.8	A large majority of students reported that Testmoz made homework more engaging than traditional paper-based assignments. The high level of agreement suggests that the platform increased students' interest and participation in learning activities.
Digital Preference	5.5	8.9	85.6	Most students preferred digital assignments over traditional paper-based homework. The high level of agreement indicates strong acceptance of digital formative assessment and suggests that students valued the flexibility and convenience offered by Testmoz.
Motivation	4.9	12.0	83.1	Most students agreed that Testmoz increased their motivation to complete assignments. The high level of agreement suggests that features such as convenience, accessibility, and immediate feedback encouraged greater participation in learning activities.
Feedback Helpfulness	5.6	8.9	85.5	Most students perceived the feedback provided by Testmoz as helpful for identifying mistakes and improving their learning. The high level of agreement highlights the value of immediate feedback in supporting self-monitoring and learning improvement.
Completion Consistency	4.4	8.8	86.7	A large majority of students reported completing assignments more consistently when using Testmoz. The high level of agreement suggests that the platform encouraged regular participation and supported students in maintaining their study routines.

Dimension	Disagree (%)	Neutral (%)	Agree (%)	Interpretation
Continuance Intention	3.8	7.7	88.5	The overwhelming majority of students expressed a desire to continue using Testmoz in future English courses. This strong level of support suggests high overall satisfaction with the platform and its perceived value for learning.

In addition, the mean scores from the five-point Likert scale further supported these positive perceptions. Continuance Intention received the highest mean score (M = 4.47), followed by Ease of Use (M = 4.43), Completion Consistency (M = 4.38), Feedback helpfulness (M = 4.30), and Motivation (M = 4.28). Across all surveyed items, over 83% of respondents selected either "Agree" or "Strongly Agree." Specifically, 70% of the total respondents opted for "Strongly Agree," while 30% selected "Agree." These high percentages reflect strong consensus and internal consistency across the dataset.

Moreover, these findings indicate that students adapted favorably to the transition from paper-based assignments to digital formative assessment. While ease of use scored the highest quantitatively, the qualitative responses suggest that the helpfulness of immediate feedback was a primary driver of student satisfaction, encouraging continued participation due to rapid access to results. However, a minor yet important trend emerged regarding a small group of students who reported lower satisfaction levels. This minority likely represents learners facing technological barriers or lower digital literacy. Although they constituted a small percentage of the sample, their feedback highlights the necessity of providing targeted support to ensure equitable participation in digital learning environments.

And some qualitative evidence open-ended responses consistently emphasized the speed and convenience of the platform: “I like Testmoz the most because it gives quick feedback on scores.”; “Online tasks can be done anywhere and quickly.”; “Reviewing the results and mistakes helps me identify my weaknesses and learn from them.” → These comments demonstrate that Testmoz improved administrative efficiency while simultaneously promoting learner autonomy by enabling students to independently monitor and evaluate their own progress.

4.2 Challenges and Areas for Improvement

Although student perceptions were generally positive, several challenges emerged from the open-ended responses.

Table 2

Main Challenges Reported by Students

Theme	Representative Comments	Educational Implication
Internet Connectivity	“The internet connection can affect the test.”	Reliable infrastructure remains essential.
Connectivity Anxiety	“I’m worried about losing the connection during the test.”	Digital assessment may introduce new forms of stress.
Limited Explanatory Feedback	“I want the app to point out mistakes more clearly.”	Automated scoring should be supplemented with pedagogical feedback.

Theme	Representative Comments	Educational Implication
Need for More Practice Activities	"I would like more practice exercises."	Students value opportunities for additional learning and revision.

The most frequently reported challenge concerned Internet stability. Students recognized the convenience of online assessment but also expressed concern about technical interruptions during quizzes. This finding suggests that successful digital transformation depends not only on software adoption but also on reliable technological infrastructure.

Another recurring theme was the demand for explanatory feedback. While students appreciated automatic scoring, several respondents expressed a desire for more detailed explanations of incorrect answers. This indicates that digital tools can improve efficiency but cannot fully replace the instructional role of teachers.

5. Discussion

The purpose of this study was to investigate the implementation of Testmoz in General English courses at Vietnam Aviation Academy (VAA) and to examine students' perceptions of the transition from paper-based to digital formative assessment. The discussion is organized around the two research questions.

5.1 Discussion of RQ1: Designing and Implementing Testmoz for Formative Assessment

The first research question explored how Testmoz could be aligned with the General English syllabus at VAA and whether its implementation improved formative assessment practices.

The findings suggest that Testmoz successfully supported the transition from traditional paper-based homework to a digital assessment environment. The platform enabled instructors to create not only quizzes that corresponded directly to vocabulary, grammar, reading, and listening objectives but also short answers for writing and collecting assignement of essay and speaking within the General English curriculum. More importantly, the automated scoring system provided immediate feedback, a feature that students consistently identified as the most valuable aspect of the platform.

Table 3

Summary of Student Perceptions

Variable	Mean Score
Ease of Use	4.43
Feedback Helpfulness	4.30
Motivation	4.28
Homework Completion	4.38
Continuance Intention	4.47

The high ratings for feedback helpfulness (M = 4.4) and continuance intention (M = 4.4) indicate that students valued the speed and convenience of digital assessment.

The ranging from 4.28 to 4.47 on a five-point Likert scale indicates extremely positive student perceptions across all measured dimensions. Continuance intention received the highest rating (M = 4.47), suggesting that students were strongly willing to accept and continue using Testmoz in future English courses. Ease of use (M = 4.43) and homework completion (M = 4.38) also received particularly favorable evaluations, showing that students found the platform accessible and believed it helped them complete assignments more conveniently.

In addition, these findings support the argument of Hattie and Timperley (2007) that timely feedback plays a critical role in improving learning outcomes. Similarly, Nicol and Macfarlane-Dick (2006) emphasized that rapid formative feedback promotes self-regulated learning by helping learners identify weaknesses and adjust their study strategies.

5.2 Discussion of RQ2: Student Perceptions, Motivation, and Learner Autonomy

The second research question mainly focused on how students perceived the transition from manual to digital assessment, particularly regarding motivation and control over their own learning because of the marks self-directed learning that accounts for up to 30% of course assessment.

Overall, student perceptions were strongly positive. Most respondents described Testmoz as easy to use, convenient, and accessible. Open-ended responses frequently referred to “quick feedback,” “easy use,” and “learning anywhere,” demonstrating that digital assessment was viewed as more flexible than traditional homework.

The findings further suggest that Testmoz enhanced learner autonomy. Students reported that immediate access to results helped them identify mistakes independently and monitor their own progress. This supports the view that formative assessment is most effective when learners actively engage with feedback rather than passively receive grades. The ability to complete quizzes outside class and instantly review performance encouraged greater ownership of learning.

Another noteworthy finding concerns mobile learning. A substantial proportion of students accessed Testmoz through smartphones, indicating that the platform functioned effectively within a Mobile-Assisted Language Learning (MALL) environment. This finding aligns with previous research suggesting that mobile accessibility increases learning flexibility and encourages continuous engagement beyond classroom boundaries. In the context of VAA, smartphone compatibility appears particularly valuable because it allows students to complete self-study activities at convenient times and locations.

Despite these positive perceptions, several challenges emerged. The most common concern involved unstable Internet connections. Students expressed anxiety about potential disconnections during quizzes, suggesting that digital assessment introduces new forms of technological stress that are absent in paper-based environments. Additionally, some students indicated that numerical scores alone were insufficient and requested clearer explanations of errors. These findings demonstrate that successful digital transformation requires attention not only to technological functionality but also to learner support and confidence.

5.3 The E-SAFE Framework for Responsible Digital Assessment

Based on the findings, this study proposes the E-SAFE framework as a practical model for implementing digital formative assessment in higher education.

Table 4

The E-SAFE Framework

Component	Description	Evidence from This Study
E – Ethical Governance	Transparency, informed consent, responsible use of learner data	Low student awareness of data privacy issues
S – Sustainable Infrastructure	Reliable and scalable digital systems	Concerns regarding unstable Internet connections
A – Academic Integrity Protection	Assessment design that discourages misconduct	Time limits and randomized questions reduced opportunities for cheating
F – Feedback Transparency	Clear and meaningful explanations of performance	Students requested more detailed explanations of mistakes
E – Equity	Inclusive access regardless of device or location	Successful use through both smartphones and laptops

The framework extends beyond technological adoption by emphasizing the human and ethical dimensions of digital assessment. In particular, the student comments concerning Internet reliability directly support the need for Sustainable Infrastructure and Equity. Institutions should therefore consider flexible assessment settings, backup submission procedures, and mobile-friendly design to reduce technology-related barriers.

Similarly, the demand for explanations rather than simple scores highlights the importance of Feedback Transparency. Automated grading should be supplemented by teacher interpretation and discussion so that students understand not only what is wrong but also why it is wrong.

The findings also raise broader ethical concerns. Although students demonstrated strong acceptance of digital assessment, only a minority reported understanding how educational platforms store and manage personal data. This suggests that digital literacy training should accompany technological implementation. Universities must ensure that students understand both the benefits and responsibilities associated with digital learning environments.

Finally, while AI-supported assessment tools create opportunities for efficiency, they should not replace professional judgment. Human oversight remains essential for interpreting learner needs, addressing misconceptions, and maintaining academic integrity. The most effective approach appears to be a balanced model in which automated systems handle routine assessment tasks while instructors focus on pedagogical guidance and learner development.

6. Conclusion

6.1 General view

This study investigated the design, implementation, and evaluation of Testmoz as a digital formative assessment tool in General English courses at Vietnam Aviation Academy (VAA). The findings provide evidence that Testmoz is a lightweight and low-cost digital assessment platform and can support the ongoing digital transformation of higher education while remaining practical for institutions with limited technological resources.

Regarding the first research question, the study strongly indicated that Testmoz could be effectively matched the General English syllabus through the development of curriculum-based quizzes targeting not only vocabulary, grammar, reading, and self-study activities but also writing and speaking. The implementation improved the speed of formative feedback and reduced instructor workload through automated grading and reporting functions. Digital transformation helps teachers spend less time on routine administrative tasks and more time on instructional support and curriculum refinement.

Regarding the second research question, students generally perceived the transition from paper-based homework to digital assessment positively. High ratings for ease of use, feedback helpfulness, homework completion, and continuance intention indicate strong acceptance of the platform. Qualitative responses further revealed that immediate feedback, convenience, and accessibility were the most valued features. Many students reported that Testmoz helped them identify mistakes, monitor learning progress, and take greater responsibility for their own learning, suggesting enhanced learner autonomy and self-regulation.

Beyond the local context, the study proposes the E-SAFE framework (Ethical Governance, Sustainable Infrastructure, Academic Integrity Protection, Feedback Transparency, and Equity) as a practical model for guiding responsible digital assessment implementation. The framework emphasizes that successful digital transformation requires not only technological adoption but also attention to ethical governance, equitable access, meaningful feedback, and sustainable infrastructure.

Overall, the study argues that digital formative assessment should be viewed not merely as a technological innovation but as a pedagogical strategy that enhances learning efficiency, supports learner autonomy, and improves instructional effectiveness. For emerging universities seeking sustainable pathways toward digital transformation, Testmoz provides an accessible example of how assessment practices can be modernized without requiring expensive or complex technological solutions.

6.2 Limitations

Here are some of the limitations that can be recognized. First, the study relied primarily on self-reported student perceptions rather than longitudinal measures of academic achievement. Second, the research was conducted at a single institution and within General English courses only. As a result, the findings may not be fully generalizable to other educational contexts, disciplines, or student populations. Third, the study focused on the implementation of one digital assessment platform and did not compare Testmoz with alternative systems. Therefore, it is not possible to determine whether the reported benefits are unique to Testmoz or representative of digital formative assessment tools more broadly. Finally, although issues related to academic integrity, digital literacy, and data privacy were discussed, these dimensions were explored through student perceptions rather than objective behavioral measures.

6.3 Contributions and Future Directions

This study contributes to the growing body of research on digital formative assessment by presenting a sustainable, low-cost, and scalable model for emerging universities. The findings demonstrate that Testmoz can enhance feedback efficiency, support learner autonomy, reduce instructor workload, and function effectively within mobile learning environments. The proposed

E-SAFE framework also provides a practical and conceptual guide for responsible digital transformation in higher education.

Building on these contributions, future research should examine the long-term impact of digital formative assessment on language learning outcomes, compare different assessment platforms, and explore integration with AI-enhanced learning systems, learning analytics, and Mobile-Assisted Language Learning (MALL). Further investigation into ethical issues, academic integrity, and self-regulated learning will help ensure that digital assessment practices remain effective, equitable, and sustainable in evolving educational contexts.

References

Bennett, S., Dawson, P., Bearman, M., Molloy, E., & Boud, D. (2017). How technology shapes assessment design: findings from a study of university teachers. British Journal of Educational Technology. https://doi.org/10.1111/bjet.12439

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge University Press.

Culatta, R. (2018, April 2). *ADDIE model.* Retrieved September 7, 2018, from http://www.instructionaldesign.org/models/addie/

Dang, T. D., Phan, T. T., Vu, T. N. Q., La, T. D., & Pham, V. K. (2024). Digital competence of lecturers and its impact on student learning value in higher education. *Heliyon, 10*(17), e37318. https://doi.org/10.1016/j.heliyon.2024.e37318

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Nhung, N. T. H., Kien, P. T., Khanh, M. Q., Tinh, T. T., & Phong, T. D. P. (2025). Digital transformation in Vietnam's education: Opportunities, challenges, and development strategies. *Multidisciplinary Reviews, 8*(9), 2025282. https://doi.org/10.31893/multirev.2025282

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Pariska, D., & Rayusda, R. A. (2026). Instructional design for meaningful english learning: bridging pedagogy, technology, and assessment. YASIN, 6(1), 989-1000. https://doi.org/10.58578/yasin.v6i1.9065

Phuong, H. Y., Nguyen, T., Pham, T., Le, T. T., Huynh, T., & Nguyen, T. (2024). Challenges of implementing outcome-based assessment in Vietnamese higher education–a qualitative inquiry. EIKI Journal of Effective Teaching Methods, 2(1). https://doi.org/10.59652/jetm.v2i1.123

Priyeti, S., Degeng, M. D. K., & Wedi, A. (2025). *Learning Management System vs conventional assessment: Reviewing the best assessment for English subjects*. *Journal La Sociale, 6*(1), 185–198. https://doi.org/10.37899/journal-la-sociale.v6i1.1800

Redecker, C., & Johannessen, Ø. (2013). Changing assessment—Towards a new assessment paradigm using ICT. *European Journal of Education, 48*(1), 79–96. https://doi.org/10.1111/ejed.12018

Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Bionote

Ms. Phan Tu Trinh is a dedicated English lecturer and researcher with over a decade of experience in language education and curriculum development. She has led and co-authored numerous research projects on CLT, MALL, discourse analysis, ESP, and flipped classroom models. Currently teaching at Vietnam Aviation, she holds a Master's in TESOL and multiple teaching certifications. Her work focuses on improving English learning outcomes through innovative methodologies and digital integration. Ms. Trinh is not only passionate about connecting Vietnamese students with global opportunities through English proficiency, but also committed to continuous learning and self-improvement.

Appendices

Student Questionnaire

(link: https://forms.gle/RSGtGtxWkZoEwJneA)

Survey Title: Student Perceptions of Using Testmoz for General English Homework

Participants: 182 students

Type: Mixed-format (Likert scale + open-ended)

Section 1: Background Information

1. Gender: ☐ Male ☐ Female ☐ Other
2. Device used to access Testmoz: ☐ Smartphone ☐ Laptop ☐ Tablet ☐ Other

---

Section 2: Student Experience with Testmoz

No.	Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I found Testmoz easy to use.	☐	☐	☐	☐	☐
2	Testmoz made English homework more engaging.	☐	☐	☐	☐	☐
3	I prefer doing homework on Testmoz over paper-based assignments.	☐	☐	☐	☐	☐
4	I completed my homework more regularly when using Testmoz.	☐	☐	☐	☐	☐
5	Receiving instant feedback helped me learn better.	☐	☐	☐	☐	☐
6	I feel more motivated to complete assignments on Testmoz.	☐	☐	☐	☐	☐
7	Testmoz helped me understand my mistakes.	☐	☐	☐	☐	☐
8	I would like to continue using Testmoz in future English classes.	☐	☐	☐	☐	☐

Section 3: Open-Ended Questions

1. What did you like most about using Testmoz for English 3 homework?
2. What difficulties did you face (if any) while using Testmoz?
3. In what ways do you think Testmoz helped improve your English learning?
4. Do you have any suggestions for improving digital homework tools like Testmoz?

122 ENGLISH TEACHER IDENTITY AND AGENCY IN A VIETNAMESE EDUCATIONAL CONTEXT: RESPONDING TO ESL-ORIENTED EXPECTATIONS

Thanh Hai Ha[1], Thi Thu Hien Nguyen [2]

Abstract

This study investigates how upper secondary English teachers in Vietnam respond to emerging ESL-oriented expectations within the country's evolving English language education context. Although recent educational directions increasingly emphasize communicative and use-oriented English teaching, limited research has explored how teachers interpret such expectations before they become fully institutionalized. The study examines how teachers perceive ESL-oriented expectations and how these perceptions shape their professional identity and agency. A qualitative research design was employed, using semi-structured interviews with eight female secondary English teachers working in urban public schools in Vietnam. The data were analyzed using reflexive thematic analysis. The findings indicate that teachers associate ESL-oriented expectations primarily with communicative teaching and authentic language use, while simultaneously experiencing uncertainty due to limited implementation guidance. Participants anticipated shifting from knowledge transmitters to facilitators of communication and learning, yet also foresaw tensions related to exam-oriented systems, large class sizes, and uneven student proficiency. Although teachers anticipated adapting professionally, their agency appeared strongly mediated by institutional and contextual constraints. The study contributes to research on teacher identity and agency by showing how professional meaning-making emerges not only through emerging ESL-oriented educational change, but also through teachers' interpretations of anticipated policy expectations. Given the context-specific nature of the study, the findings should be understood as reflecting the experiences of female English teachers working in urban public secondary schools rather than Vietnamese teachers more broadly.

***Keywords*:** teacher identity, professional agency, ESL-oriented expectations, educational change, Vietnam

1. Introduction

English language education has undergone significant transformation in many contexts, particularly through increasing emphasis on communicative and use-oriented approaches to English learning. In Vietnam, this shift is reflected in ongoing efforts to move beyond an exam-oriented EFL model toward approaches that emphasize real-life language use and communication. However, such changes are often articulated as broad orientations rather than fully specified and systematically implemented policies, leaving teachers to interpret and respond to evolving expectations in their own ways.

In this study, the term ESL-oriented does not suggest that Vietnam is becoming a traditional ESL society in which English functions as a dominant societal language. Rather, it refers to

[1] *Corresponding author, Doctor, Department of Foreign Languages, Quy Nhon University, Gia Lai, Vietnam; hathanhhai@qnu.edu.vn*
[2] *Assoc. Prof. Dr, Department of Foreign Languages, Quy Nhon University, Gia Lai, Vietnam.*

emerging expectations for English to serve increasingly as a medium for communication, participation, and learning within educational contexts. The term therefore captures a pedagogical and functional orientation toward English use rather than a complete sociolinguistic transition.

Within this context, research on language teacher identity provides a useful lens for understanding how teachers engage with change. Teacher identity has been widely conceptualized as dynamic, socially constructed, and context-dependent (Beijaard et al., 2004; Varghese et al., 2005). Rather than being fixed, identity is continuously negotiated through teachers' interactions with institutional discourses, professional experiences, and sociocultural expectations. From this perspective, changes in educational orientations - such as a shift toward communicative or ESL-oriented approaches - are likely to reshape how teachers understand their roles and responsibilities.

While teacher identity helps explain how teachers understand their professional roles, the concept of teacher agency provides insight into how they respond to educational change in practice. From an ecological perspective, agency is not an inherent capacity but an achievement emerging from the interaction between individuals and their contexts (Priestley et al., 2015; Biesta & Tedder, 2007). Teachers are therefore not passive implementers of policy, but active agents who interpret, adapt, and negotiate pedagogical expectations within the constraints and affordances of their working environments. Understanding teachers' anticipatory responses to educational reform is important because teachers' interpretations often shape how policy directions are eventually enacted in practice. Examining teachers' perceptions at this emerging stage therefore provides insight not only into possible implementation challenges, but also into how educational change is interpreted and imagined before formal institutionalization occurs.

At the same time, a substantial body of research has documented tensions between policy expectations and classroom realities. Teachers' beliefs, prior experiences, and contextual conditions often shape how educational reforms are interpreted and enacted, leading to discrepancies between intended policy goals and actual classroom practices (Borg, 2006). From a policy enactment perspective, educational change is not implemented uniformly but interpreted and negotiated differently across local contexts (Ball et al., 2012). Such tensions become particularly visible in contexts where teachers are expected to adopt new professional roles without sufficient structural support, institutional alignment, or clear implementation guidance.

In Vietnam, recent studies have highlighted similar challenges, including large class sizes, exam-oriented assessment systems, and uneven student proficiency levels, all of which complicate the implementation of communicative approaches (Do, 2022; Vu & Burns, 2014). While these studies provide valuable insights into the difficulties of reform implementation, they have largely focused on enacted practices rather than on how teachers perceive and respond to emerging expectations before such expectations become clearly institutionalized.

This points to an important gap in the literature. While teacher identity and agency have been extensively studied, less attention has been paid to how these constructs are shaped in contexts where policy directions remain evolving and not yet fully institutionalized. In such contexts, teachers are not responding to clearly defined requirements, but to anticipated and interpreted expectations, which may influence how they understand their roles and prepare for change.

Addressing this gap, the present study examines how English teachers in a Vietnamese educational context perceive, interpret, and respond to emerging ESL-oriented expectations. By focusing on teachers' perceptions, anticipated challenges, and projected responses, the study seeks to contribute to a more nuanced understanding of educational change as a process shaped not only through implementation, but also through interpretation and anticipation.

To address these issues, the study investigates how English teachers perceive and respond to emerging ESL-oriented expectations, with particular attention to professional identity, anticipated challenges, and professional agency. With this aim, the study addresses the following research questions

RQ1. How do English teachers in Vietnam perceive their changing roles in response to emerging ESL-oriented expectations?

RQ2. What tensions and challenges do they anticipate in responding to ESL-oriented expectations?

RQ3. How do they anticipate exercising professional agency in negotiating policy expectations and classroom realities?

While previous studies have primarily examined teachers' experiences after communicative reforms have been implemented, considerably less attention has been paid to how teachers interpret and make professional meaning from anticipated ESL-oriented expectations before such reforms become fully institutionalized. This study addresses this earlier stage of educational change.

2. Literature Review

2.1. Vietnam's shift toward an ESL-oriented educational context

English language education in Vietnam has undergone substantial transformation in recent years as the government increasingly positions English as a strategic resource for international integration, economic development, and educational modernization. This shift is reflected in recent policy documents such as Kết luận số 91-KL/TW (Bộ Chính trị, 2024) and Quyết định số 2371/QĐ-TTg (Thủ tướng Chính phủ, 2025), which articulate the national vision of gradually bringing English into schools as a second language. These policy directions indicate a movement away from a traditional exam-oriented EFL model toward a more communicative and use-oriented approach to English education.

Within a conventional EFL framework, English is generally treated as a discrete academic subject emphasizing grammar, vocabulary, and examination performance. By contrast, an ESL-oriented perspective tends to position English more functionally, emphasizing communication, participation, and practical language use within educational and social contexts. However, Vietnam cannot yet be considered a fully developed ESL society in the traditional sense. Rather, the country appears to be moving toward an ESL-oriented educational context, where English is increasingly expected to function as a communicative and educational resource within schooling.

Recent scholarships suggest that this transition involves more than curriculum reform alone. Ngo (2023), for example, argues that English in Vietnam is increasingly associated with employability, mobility, and globalization, yet the implementation of communicative teaching remains uneven across educational contexts. This indicates that policy aspirations alone are

insufficient to transform classroom practice unless accompanied by pedagogical, institutional, and professional change.

Educational reform does not occur uniformly across contexts. Policy enactment perspectives emphasize that schools and teachers actively interpret and negotiate policy within local institutional realities rather than mechanically implementing official directives (Ball et al., 2012; Braun et al., 2011). This is particularly relevant in Vietnam, where differences in resources, assessment systems, and institutional conditions may shape how ESL-oriented expectations are understood and enacted.

Such changes inevitably place English teachers at the center of educational transition. As English becomes increasingly positioned as a resource for communication and participation rather than solely as an object of study, teachers may be expected to assume new pedagogical and professional roles. Understanding how teachers perceive and respond to these emerging expectations therefore becomes an important area of inquiry.

2.2. Teacher identity and the redefinition of professional roles

The concept of teacher identity has become increasingly important in research on educational reform because it helps explain how teachers understand themselves and their professional roles within changing institutional contexts. Beauchamp and Thomas (2009) conceptualize teacher identity as a dynamic and complex construct shaped by personal experience, professional interpretation, and contextual influences rather than fixed individual characteristics. From this perspective, educational change often requires teachers not only to adopt new practices but also to renegotiate their professional sense of self.

Research in language teacher identity has similarly emphasized the socially situated and continuously evolving nature of identity construction. Barkhuizen (2016, 2017) argues that language teacher identity is constructed and reconstructed through interaction with institutional discourses, policy expectations, classroom experiences, and sociocultural contexts. Language teachers are therefore not passive recipients of reform; rather, they actively interpret and negotiate what it means to teach language within specific educational environments.

In contexts of educational transition, identity is often future-oriented. Ruohotie-Lyhty (2018) suggests that teachers construct imagined professional selves in response to anticipated expectations and projected professional demands, while Zembylas (2003) highlights the emotional dimensions accompanying identity reconstruction during periods of change. These perspectives are particularly relevant in Vietnam's emerging ESL-oriented context, where teachers may increasingly be expected to move beyond traditional knowledge-transmission roles and adopt more communicative and facilitative professional identities.

In the present study, teacher identity serves as a lens for examining how teachers perceive and reconstruct who they are becoming in response to emerging ESL-oriented expectations.

2.3. Professional agency and educational change

Professional agency provides insight into how teachers respond to educational change within the opportunities and constraints of their professional contexts. Priestley, Biesta, and

Robinson (2015) conceptualize teacher agency from an ecological perspective, arguing that agency is not an inherent individual trait but an achievement emerging from the interaction between teachers' prior experiences, present working conditions, and future aspirations.

From this perspective, teachers' responses to reform are shaped not only by their motivation or beliefs, but also by contextual conditions such as curriculum demands, institutional culture, assessment systems, professional support, and available resources. Agency is therefore relational and contextually mediated rather than purely individual.

Existing studies in Vietnam have also documented tensions associated with English-related educational reforms, including language proficiency concerns, pedagogical pressures, unequal learning conditions, and contextual constraints (Do, 2022; Vu & Burns, 2014; Vo et al., 2022). These studies suggest that teachers' responses to communicative and English-related expectations are strongly shaped by institutional realities and local teaching conditions. More recently, Dang (2024) highlights the role of transformative agency among Vietnamese teachers, showing that teachers actively reinterpret and negotiate reform expectations within sociocultural and institutional constraints.

Taken together, these studies suggest that teachers' responses to ESL-oriented expectations are likely to involve ongoing negotiation between emerging professional demands and existing structural realities. However, limited research has explored how teachers imagine responding to such expectations before they become fully institutionalized. In the present study, professional agency provides a lens for examining teachers' projected actions, adaptation strategies, and responses to anticipated educational change.

2.4. Theoretical framework

The literature reviewed above highlights three complementary perspectives relevant to understanding teachers' responses to emerging ESL-oriented expectations. Teacher identity explains how teachers perceive and reconstruct their professional roles. Professional agency explains how teachers imagine responding to educational change within contextual opportunities and constraints. Policy enactment contributes a broader perspective by explaining how reform expectations are interpreted, negotiated, and translated differently across local educational contexts.

This study is grounded in a socio-constructivist perspective of teacher identity, an ecological perspective of professional agency (Priestley et al., 2015), and the concept of policy enactment (Ball et al., 2012; Braun et al., 2011). Together, these perspectives conceptualize teachers' responses to ESL-oriented expectations as a dynamic interaction among evolving professional identity, contextually mediated professional agency, and institutional conditions shaping educational change.

This integrated framework provides the conceptual basis for examining how teachers make sense of emerging ESL-oriented expectations and project their professional responses in contexts where reform directions remain evolving and not yet fully institutionalized.

The framework also informed the design of the interview protocol and the analytical process. Teacher identity guided questions related to perceived role changes, professional responsibilities, and imagined future selves. Professional agency informed questions concerning projected actions, adaptation strategies, and the negotiation of contextual constraints. Policy

enactment informed questions exploring how teachers interpreted emerging and sometimes unclear reform expectations. During analysis, these concepts served as sensitizing lenses for coding and theme development, while still allowing new themes to emerge inductively from the data.

3. Methodology

3.1. Research design

This study adopts a qualitative research design to explore how upper secondary English teachers in Vietnam perceive and respond to emerging ESL-oriented expectations. A qualitative approach was considered appropriate because the study focuses on teachers' interpretations, perceptions, and projected responses to educational change within their specific professional contexts.

The study is informed by a socio-constructivist perspective, which views teacher identity and agency as socially constructed and context-dependent. Since ESL-oriented expectations in Vietnam have not yet been fully institutionalized, the study focuses on how teachers interpret and anticipate these emerging expectations rather than on actual policy implementation.

More specifically, the study may be characterized as an interpretive qualitative interview study informed by teacher identity research, ecological perspectives on professional agency, and policy enactment theory.

3.2. Research participants and participant coding

The participants included eight female English teachers working in urban public upper secondary schools. Participants ranged in age from 27 to 49 years old and had between 6 and 25 years of teaching experience. All participants held bachelor's degrees in English Language Teacher Education, and five had completed master's degrees in TESOL. Purposive sampling was employed to recruit information-rich participants who were likely to provide meaningful insights into emerging ESL-oriented expectations. Participants were selected because they were experienced English teachers working in public upper secondary schools where communicative teaching reforms and English language policy developments were increasingly relevant. Variation in age, teaching experience, and academic qualifications was also considered during recruitment to capture a range of professional perspectives.

The participants worked within the context of Vietnam's 2018 General Education English Curriculum, which officially emphasizes communicative competence and the integrated development of four language skills. However, English teaching and learning at the secondary level remain strongly influenced by examination-oriented assessment systems. Although current textbooks include communicative and four-skill activities, classroom instruction often focuses on language knowledge and test-oriented content related to school examinations, high school graduation examinations, and university entrance assessments. Major examinations continue to rely heavily on multiple-choice formats, while listening, speaking, and extended writing are rarely formally assessed. At the upper secondary level, English is generally taught approximately three periods per week under the national curriculum. As a result, participants worked within educational contexts characterized by tensions between communicative curriculum goals and exam-oriented classroom realities.

To ensure confidentiality, pseudonymized participant codes were used throughout transcription, coding, and reporting. Each participant was assigned a numerical code (e.g., T1, T2, T3), in which T refers to Teacher.

Although the sample provided rich professional perspectives, it represents a relatively specific group of participants. All teachers were female and worked in urban public upper secondary schools. Consequently, the findings should be interpreted within this particular educational context and may not fully reflect the experiences of male teachers, teachers working in rural or disadvantaged areas, private-school teachers, or teachers operating under different institutional conditions. The purpose of the study, however, was not statistical generalization but in-depth exploration of teachers' professional meaning-making in a specific educational setting.

3.3. Data collection

Data were collected through semi-structured individual interviews conducted over approximately three weeks in March 2026. Each interview lasted between 40 and 60 minutes and was conducted primarily in Vietnamese to allow participants to express their ideas comfortably and in depth.

The interview questions were developed based on the study's research questions and theoretical framework, particularly the concepts of teacher identity, professional agency, and policy enactment. Because ESL-oriented expectations remain emerging rather than fully implemented, many questions were framed using future-oriented language to explore teachers' anticipations, imagined roles, and projected responses.

The interview protocol focused on three main areas aligned with the research questions: (1) teachers' perceptions of changing professional roles under emerging ESL-oriented expectations; (2) anticipated tensions and challenges related to classroom realities and institutional conditions; and (3) teachers' projected professional agency in responding to evolving educational expectations. Prior to data collection, the interview protocol was reviewed by two colleagues with expertise in English language education and qualitative research. Minor revisions were made to improve question clarity, sequencing, and alignment with the research questions and theoretical framework. Sample interview questions are provided in Appendix A.

Although the interview questions were organized around these broad areas, participants were also encouraged to elaborate on related experiences, emotions, concerns, and professional expectations that emerged during the interviews. All interviews were audio-recorded with participants' consent and later transcribed for analysis.

3.4. Data analysis

The data were analyzed using reflexive thematic analysis following Braun and Clarke (2006, 2021). Interview transcripts were read multiple times to achieve familiarity with the data before initial codes were generated.

The coding process combined deductive and inductive approaches. Initial coding categories were informed by the study's research questions and theoretical framework, particularly the concepts of teacher identity, professional agency, and policy enactment. At the same time, additional patterns related to uncertainty, ambiguity, emotional pressure, and imagined professional roles emerged inductively during the analysis.

Related codes were subsequently grouped into broader themes representing shared patterns across participants' responses. The final themes focused on teachers' perceived role changes under emerging ESL-oriented expectations; anticipated tensions between educational expectations and contextual realities; and projected professional agency in responding to emerging educational change.

The analysis adopted an interpretive rather than purely descriptive orientation, focusing not only on what teachers said but also on how they constructed meaning around anticipated educational change, uncertainty, and professional transformation.

To enhance the trustworthiness of the analysis, the researchers engaged in repeated transcript reading, iterative coding, and ongoing analytical discussions throughout the study. Reflexive notes were maintained to document emerging interpretations, coding decisions, and potential researcher assumptions. Themes were reviewed repeatedly against the complete dataset to ensure coherence, consistency, and adequate representation of participants' perspectives. Consistent with reflexive thematic analysis, the emphasis was placed on analytic transparency and reflexivity rather than inter-rater reliability (Braun & Clarke, 2021).

3.5. Ethical considerations

All participants were informed about the purpose of the study before the interviews were conducted. Participation was voluntary, and participants had the right to withdraw from the study at any time. To protect confidentiality, all names and identifying information were anonymized during transcription and reporting.

4. Results

4.1. Perceived role changes under emerging ESL-oriented expectations

Six of the eight participants associated emerging ESL-oriented expectations with a stronger emphasis on communicative competence and authentic language use within educational settings. Teachers perceived that future expectations would extend beyond traditional exam-oriented instruction, requiring them to support students' ability to communicate effectively both inside and outside the classroom. As one participant explained:

"Chắc chắn sẽ dạy cho sinh viên giao tiếp nhiều hơn… làm sao cho học sinh, sinh viên có thể giao tiếp được trong phạm vi nhà trường…" (T3)

Three participants (T5, T7, T8) also anticipated broader professional responsibilities extending beyond language teaching itself. They described English teachers as potentially becoming a central force in supporting communication across the school environment, including assisting colleagues from other subject areas in using English more confidently.

These perceptions suggest that teachers anticipated ESL-oriented expectations as requiring a shift from knowledge transmission toward more facilitative, communicative, and supportive professional roles. Participants envisioned themselves not only as instructors, but also as facilitators, encouragers, consultants, and creators of English-speaking environments:

"Có lẽ không chỉ là một người giảng bài, mà có thể đóng nhiều vai facilitator, encourager, cooperator…" (T1)

"Tôi nghĩ mình sẽ là một cooperator cùng tham gia các hoạt động ngôn ngữ với các em" (T2)

In addition, teachers associated these changing roles with increased interpersonal, emotional, and professional demands. Participants like T1, T4, T6, and T7 described the future English teacher as needing to become more flexible, communicative, active, and professionally responsible. Such responses indicate an emerging reconstruction of teacher identity shaped by anticipated expectations regarding communication-oriented teaching and authentic language use.

However, despite this general understanding, three participants (T4, T7, and T8) consistently expressed uncertainty regarding how such expectations would eventually be implemented in practice. T4, for example, described current directions as "mơ hồ" (unclear) and difficult to predict:

"Tôi thấy khá mơ hồ… nhưng chắc sẽ khó khăn." (T4)

This perceived ambiguity appeared to shape how teachers imagined their future professional roles. While most participants viewed communicative expectations positively, a few teachers (T1, T3 and T4) also expressed skepticism regarding the feasibility of such expectations within Vietnam's highly exam-oriented educational system. In addition, some participants expressed anxiety about whether they would be capable of fulfilling these expanding professional responsibilities:

"Nhiều responsibilities hơn nên lo lắng không biết có làm tròn không." (T2)

These findings suggest that teachers are actively constructing future-oriented professional identities based not on clearly institutionalized policy requirements, but on their interpretations of evolving and anticipated educational expectations.

4.2. Anticipated tensions between ESL-oriented expectations and contextual realities

Nearly all participants (seven of the eight teachers) anticipated significant challenges in aligning ESL-oriented expectations with existing classroom and institutional conditions. A key concern was the mismatch between communicative teaching approaches and structural constraints such as large class sizes, exam-oriented systems, and uneven student proficiency levels:

"*Lớp hiện tại rất đông không phù hợp để dạy giao tiếp… thi cử còn nặng nề… đặc biệt ở các lớp cuối cấp." (T5)*

Teachers also pointed to broader systemic issues, including disparities across regions and limitations in teachers' own language proficiency:

"*Một số giáo viên chưa đạt chuẩn để giao tiếp bằng tiếng Anh lưu loát… trình độ học sinh cũng còn thấp và chênh lệch giữa các vùng miền." (T7)*

These anticipated tensions reflect a perceived gap between policy aspirations and classroom realities. Teachers expressed concern that without structural changes, the expected shift toward ESL-oriented teaching may be difficult to achieve:

"*Không có môi trường giao tiếp bên ngoài lớp học… thi cử chưa thay đổi thì chưa đạt kỳ vọng được." (T8)*

In addition, economic and institutional factors were also identified as potential barriers:

"Lương giáo viên quá thấp nữa." (T2)

Such responses indicate that teachers are not only aware of pedagogical challenges but also critically evaluating the feasibility of reform within existing systemic conditions.

Moreover, several participants (T2, T5, T7, and T8) anticipated significant pressure associated with these changes:

"*Chắc chắn là áp lực rồi… không chỉ về cách dạy mà còn nâng cao khả năng tiếng Anh nữa." (T8)*

This highlights that the anticipated transition involves not only instructional adjustments but also substantial demands for professional development, often without clear support structures.

4.3. Professional agency under emerging expectations

Participants projected a forward-looking sense of professional agency, expressing intentions to adapt their teaching practices in response to anticipated ESL-oriented expectations. A key strategy identified was self-directed professional development, particularly in improving their own English proficiency:

"Tôi nghĩ đầu tiên mình phải tự nâng cao năng lực tiếng Anh của mình…" (T1)

In addition, some other teachers envisioned creating more opportunities for authentic language use within school settings, such as increasing communicative activities and fostering English-speaking environments:

"*Có thể tạo sân chơi cho các em nhiều hơn… nếu sân trường là một môi trường bắt buộc giao tiếp bằng tiếng Anh… thì sẽ từ từ improve mọi thứ." (T3)*

These responses suggest that teachers anticipated taking proactive steps to align their practices with anticipated expectations, particularly in terms of pedagogy and learning environment.

However, this sense of agency was not entirely autonomous. Most participants (six of the eight teachers) emphasized the need for formal training and theoretical grounding, particularly in areas such as second language acquisition:

"*Chúng tôi cần được đào tạo về second language acquisition… đầu tiên cần thay đổi trọng tâm bài học sang hướng giao tiếp." (T6)*

This suggests that while teachers anticipated being open to change, their capacity to act is closely tied to access to professional development and institutional support.

Moreover, teachers' projected agency appeared to be significantly constrained by systemic factors, particularly the persistence of exam-oriented assessment systems. One participant highlighted a reluctance to implement changes without clear alignment at the policy level:

"*Tôi sẽ chờ hướng dẫn cụ thể… nếu tôi thay đổi cách dạy, sợ rằng khi thi cử các em không đạt yêu cầu…" (T5)*

This reflects a form of projected and conditional agency, in which teachers' willingness to innovate is mediated by external structures such as curriculum and assessment. Rather than

anticipating independent pedagogical change, participants projected themselves as responding to top-down guidance and system-wide reform initiatives.

5. Discussion

This study set out to examine how English teachers in a Vietnamese educational context perceive and respond to emerging ESL-oriented expectations, with a particular focus on teacher identity and professional agency. The findings reveal a complex and dynamic process in which teachers actively interpret, anticipate, and negotiate their roles under conditions where expectations remain evolving rather than fully defined.

5.1. Teacher identity as anticipatory and negotiated

The findings suggest that teachers' professional identities are dynamic, future-oriented, and continuously negotiated in response to emerging ESL-oriented expectations. Consistent with Beauchamp and Thomas's (2009) view of teacher identity as socially constructed and evolving, participants in this study did not describe their roles as fixed or stable. Instead, they imagined themselves adopting more facilitative, communicative, and relational roles within increasingly communication-oriented educational environments.

These findings also support Barkhuizen's (2017) argument that language teacher identity is shaped through interaction with institutional and sociocultural discourses. However, unlike many previous studies focusing on enacted reform, the present study demonstrates that identity reconstruction may begin even before educational expectations become fully institutionalized. Teachers were not responding to clearly operationalized policy requirements, but rather to evolving and interpreted expectations regarding future educational change.

In line with Ruohotie-Lyhty's (2018) notion of imagined professional selves, teachers projected themselves into future-oriented identities associated with communication, flexibility, emotional engagement, and expanded professional responsibility. Their responses suggest that teacher identity is shaped not only by present classroom realities but also by teachers' interpretations of anticipated professional demands.

At the same time, this process of identity negotiation was accompanied by considerable uncertainty and emotional tension. Participants consistently described ESL-oriented expectations as unclear and difficult to predict, yet they still attempted to construct meaningful understandings of their future roles. Consistent with Zembylas (2003), concerns about competence, preparedness, and expanding responsibilities highlight the emotional dimensions of professional identity during periods of educational change.

Taken together, these findings suggest that professional identity reconstruction may begin before reform implementation occurs, as teachers actively construct meanings around anticipated rather than enacted educational expectations.

5.2. Anticipated tensions and policy ambiguity

The findings indicate that teachers anticipate significant tensions between emerging ESL-oriented expectations and existing classroom realities. Consistent with previous research on policy- practice gaps (Borg, 2006), participants identified multiple structural constraints that may hinder the realization of communicative and use-oriented teaching approaches, including large class sizes, exam-oriented assessment systems, uneven student proficiency, limited English-

speaking environments, and regional disparities. These findings also align with previous studies highlighting contextual barriers affecting English language education reform in Vietnam (Do, 2022; Vu & Burns, 2014; Vo et al., 2022).

Importantly, these tensions were largely anticipatory rather than based on direct implementation experience. Nevertheless, they already shaped how teachers interpreted emerging expectations and imagined their future practices. This suggests that the effects of educational reform may emerge not only through implementation, but also through anticipation and interpretation.

The findings also support Ball et al.'s (2012) policy enactment perspective, which argues that educational reform is interpreted and negotiated differently across local contexts rather than mechanically implemented. Even in the absence of clearly operationalized policy directives, teachers in this study were already constructing localized meanings around ESL-oriented expectations based on their professional experiences and contextual realities.

It is important to note, however, that these interpretations emerged from experienced teachers working in urban public secondary schools. Therefore, the findings should be interpreted within this particular context, as teachers in other educational settings may experience ESL-oriented expectations differently.

Overall, the findings suggest that anticipated reform may generate forms of uncertainty and tension long before institutional implementation occurs, highlighting the importance of examining educational change as an interpretive and anticipatory process rather than solely an implementation-driven phenomenon.

5.3. Professional agency under emerging reform expectations

The findings further suggest that teachers projected a forward-looking and contextually mediated form of professional agency in responding to emerging ESL-oriented expectations. Participants anticipated improving their English proficiency, incorporating more communicative activities, and creating environments that encourage authentic language use within schools. These responses indicate that teachers were actively imagining and preparing for possible future professional demands rather than passively awaiting policy implementation.

Such findings resonate with the ecological perspective on teacher agency proposed by Priestley et al. (2015), which conceptualizes agency as emerging from the interaction between individual capacities and contextual conditions. Participants' projected responses were shaped not only by motivation and professional commitment, but also by institutional structures, assessment systems, workload, and access to professional support.

At the same time, participants' projected agency appeared conditional rather than fully autonomous. Teachers consistently emphasized the need for clearer policy guidance, professional development opportunities, and systemic alignment before implementing substantial pedagogical change. Several participants expressed hesitation about independently adopting communicative approaches if assessment systems remained exam-oriented. This reflects a form of constrained agency in which teachers' willingness to innovate is mediated by broader structural realities.

The findings also align with Dang's (2024) discussion of transformative agency in Vietnam, which highlights teachers' active attempts to reinterpret and negotiate educational

expectations within sociocultural and institutional constraints. However, unlike teachers responding to enacted reforms, participants in the present study were responding primarily to anticipated and evolving expectations. This suggests that professional agency may begin to emerge before educational reforms become fully institutionalized.

Together, these findings suggest that teachers' projected responses to ESL-oriented expectations involve both identity reconstruction and contextually mediated forms of professional agency, highlighting the importance of examining educational change before formal implementation occurs.

6. Conclusion

This study examined how English teachers in a Vietnamese educational context perceive and respond to emerging ESL-oriented expectations, with particular attention to teacher identity, anticipated tensions, and professional agency. The findings revealed that teachers generally associated ESL-oriented expectations with communicative language use, authentic interaction, and expanded professional responsibilities beyond traditional exam-oriented teaching. In response to these anticipated expectations, participants projected more facilitative, relational, and communication-oriented professional identities, suggesting that teacher identity reconstruction may begin even before reforms become fully institutionalized.

At the same time, the study found that teachers anticipated significant tensions between emerging expectations and existing educational realities, including exam-oriented assessment systems, large class sizes, uneven student proficiency, and limited institutional support. Despite these challenges, participants expressed a forward-looking sense of professional agency through their willingness to improve their own professional competence and adapt their teaching practices. However, such agency appeared strongly mediated by structural and contextual conditions, particularly policy coherence and systemic alignment.

Overall, the study suggests that educational change should be understood not only as a process of policy implementation, but also as a process of interpretation, anticipation, and professional meaning-making. The findings also imply that supporting teachers in emerging reform contexts may require not only individual professional development, but also clearer policy guidance, greater alignment between assessment systems and communicative teaching goals, and stronger institutional support for teachers navigating educational change.

References

Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge. https://doi.org/10.4324/9780203153185

Barkhuizen, G. (Ed.). (2017). *Reflections on language teacher identity research*. Routledge. https://doi.org/10.4324/9781315643465

Barkhuizen, G. (2016). Language teacher identity research: An introduction. In G. Barkhuizen (Ed.), *Reflections on language teacher identity research* (pp. 1–11). Routledge. https://doi.org/10.4324/9781315643465-5

Beauchamp, C., & Thomas, L. (2009). Understanding teacher identity: An overview of issues in the literature and implications for teacher education. *Cambridge Journal of Education*, *39*(2), 175–189. https://doi.org/10.1080/03057640902902252

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(2), 107–128. https://doi.org/10.1016/j.tate.2003.07.001

Biesta, G., & Tedder, M. (2007). Agency and learning in the lifecourse: Towards an ecological perspective. *Studies in the Education of Adults, 39*(2), 132–149. https://doi.org/10.1080/02660830.2007.11661545

Borg, S. (2006). *Teacher cognition and language education*. Continuum.

Bộ Chính trị. (2024, August 12). Kết luận số 91-KL/TW về tiếp tục thực hiện Nghị quyết số 29-NQ/TW về đổi mới căn bản, toàn diện giáo dục và đào tạo [Conclusion No. 91-KL/TW on continuing the implementation of Resolution No. 29-NQ/TW on fundamental and comprehensive educational reform].

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage.

Braun, A., Ball, S. J., Maguire, M., & Hoskins, K. (2011). Taking context seriously: Towards explaining policy enactments in the secondary school. *Discourse: Studies in the Cultural Politics of Education*, *32*(4), 585–596. https://doi.org/10.1080/01596306.2011.601555

Dang, T. K. A. (2024). Teachers' transformative agency in English-medium instruction in higher education in Vietnam: A cultural-historical theory perspective. *Journal of Multilingual and Multicultural Development*. Advance online publication. https://doi.org/10.1080/01434632.2024.2349803

Do, T. T. T. (2022). Primary English language education policy in Vietnam's disadvantaged areas: Implementation barriers. *Education Sciences*, *12*(7), Article 445. https://doi.org/10.3390/educsci12070445

Ngo, M. T. (2023). Current English education in Vietnam: Policy, practices, and challenges. In Tran, L. H. N., Tran, L. T., & Ngo, M. T. (Eds.), *English language education and employability in Vietnam* (pp. 47–61). Springer. https://doi.org/10.1007/978-981-99-4338-8_3

Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury Academic.

Ruohotie-Lyhty, M. (2018). Identity-agency in progress: Teachers authoring their identities. In P. A. Schutz et al. (Eds.), *Research on teacher identity* (pp. 25–36). Springer.

Thủ tướng Chính phủ. (2025, October 27). Quyết định số 2371/QĐ-TTg phê duyệt Đề án "Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học giai đoạn 2025–2035, tầm nhìn đến năm 2045" [Decision No. 2371/QĐ-TTg approving the Project "Making English a Second Language in Schools for the Period 2025–2035, with a Vision to 2045"].

Varghese, M., Morgan, B., Johnston, B., & Johnson, K. A. (2005). Theorizing language teacher identity: Three perspectives and beyond. *Journal of Language, Identity, and Education, 4*(1), 21–44. https://doi.org/10.1207/s15327701jlie0401_2

Vo, T. D., Gleeson, M., & Starkey, L. (2022). The glocalisation of English-medium instruction examined through the ROAD-MAPPING framework: A case study of teachers and students in a Vietnamese university. *System*, *108*, Article 102856. https://doi.org/10.1016/j.system.2022.102856

Vu, N. T. T., & Burns, A. (2014). English as a medium of instruction: Challenges for Vietnamese tertiary lecturers. *The Journal of Asia TEFL*, *11*(3), 1–31.

Zembylas, M. (2003). Emotions and teacher identity: A poststructural perspective. *Teachers and Teaching*, *9*(3), 213–238. https://doi.org/10.1080/13540600309378

Bionote

Dr. Ha Thanh Hai has been an English lecturer at Quy Nhon University for more than 33 years. He earned his Master's degree in Educational Studies from The University of Adelaide and a PhD in Contrastive Linguistics. Throughout his academic career, he has published extensively in national and international scholarly journals. His research interests include academic writing, cognitive linguistics, second language acquisition, discourse analysis, and English language education. Email: hathanhhai@qnu.edu.vn.

Assoc. Prof. Dr. Nguyen Thi Thu Hien has been teaching English at Quy Nhon University for over 34 years. She holds a Master's degree in Applied Linguistics and a PhD in Contrastive Linguistics. She has authored numerous publications in professional and peer-reviewed academic journals. Her research interests include English language teaching methodology, academic writing, academic presentation, ESP, multimodality, and discourse analysis. Email: nguyenthithuhien@qnu.edu.vn.

Appendix A. Sample Interview Questions

1. How do you understand the emerging shift toward a more ESL-oriented approach to English education in Vietnam?

2. How do you think your role as an English teacher may change under these emerging expectations?

3. What challenges or tensions do you anticipate when responding to such expectations?

4. What actions might you take to adapt to these changes?

5. What forms of support would help you respond effectively to emerging ESL-oriented expectations?

#123 THE PERCEIVED USE OF METACOGNITIVE ONLINE READING STRATEGIES AMONG STUDENTS AT AN EMI UNIVERSITY IN VIETNAM

Thinh Tien Vu[1], Thi Thuy Dung Dang[2]

Abstract

In the age of technology, reading online materials has become a fundamental skill for acquiring, storing, and retrieving information, particularly as higher education institutions increasingly rely on digital resources. Despite the growing popularity of online reading, the perception of Metacognitive Online Reading Strategies (MORS) among students, especially in English-medium instruction (EMI) settings in Vietnam, remains insufficiently examined. This quantitative study addressed this gap by investigating how students at an EMI university in Vietnam perceived the use of MORS. Participants were 124 students, evenly divided between English and non-English majors. Data were collected via a survey, focusing on three sub-categories of MORS: global, problem-solving, and support strategies. Descriptive analyses indicated that problem-solving strategies were most frequently employed, followed by global and support strategies. These findings contribute to a foundation for further investigation into the application of MORS in digitally mediated EMI higher education in Vietnam.

***Keywords*:** EMI, digital reading, metacognitive online reading strategies, MORS

1. Introduction

Among the four skills in learning English, reading is considered the most important because it involves acquiring, storing, and retrieving information from texts (Amer et al., 2010; Pardede, 2022). Pardede (2022) stated three positive consequences of reading: increased input from daily interactions, enhanced learners' competence in other language-related skills, and students' personal development. Given the importance of reading skills, effectively applying reading strategies is vital for enhancing reading comprehension (Huang et al.,2009). Much research indicates that successful readers use a wider range of reading strategies more frequently than less successful readers do. (Carrell et al., 1989; Wang, 2016; Kung, 2019; Aydinbek, 2021).

Since the COVID-19 outbreak in 2020, advances in information and communication technology have made the internet a significant factor in many aspects of education, particularly in foreign language learning (Pardede, 2022; Putera & Anggrainy, 2024). In terms of technology use in the educational environment, its integration has led students to become more interactive and active rather than passive and reactive (Raja & Nagasubramani, 2018). The benefits of emerging technologies for education include enhanced teaching and learning, globalisation, removal of geographical limitations, scalability, easy-to-update learning materials, access to materials, self-

[1] *Corresponding author, PhD candidate, School of Languages, International University - VNU HCM, HCMC, Vietnam; vtthinh@hcmiu.edu.vn*
[2] *BA, School of Languages, International University - VNU HCM, HCMC, Vietnam*

management of learning location, and various designs and delivery methods for mobile technology (Ally, 2013; Raja & Nagasubramani, 2018). Thus, the use of electronic learning materials has become the norm in many countries (Parker & Lenhart, 2011). Parker and Lenhart (2011) stated that most textbooks may shift from offline materials to online materials within 10 years, so students need to apply online reading skills to obtain information from online resources. However, the transfer of reading skills from offline to online texts may be similar, but they differ significantly in purpose.

In the context of EMI universities, reading academic texts is an important skill with university students. In particular, students are often required to read textbooks and learning materials which are not available in their mother tongue (Phuong & Nguyen, 2019). These materials are usually available online rather than offline. Furthermore, EMI university students find it significantly challenging to comprehend English texts due to limited vocabulary and academic knowledge (Phuong & Nguyen, 2019). Therefore, applying online reading strategies is remarkably essential in the EMI context.

While much research focuses on metacognitive offline reading strategies, metacognitive online reading strategies (MORS) have received limited attention. Additionally, existing literature has paid little attention to the use of online reading strategies in the English-as-a-Medium-of-Instruction (EMI) context. Given the above-mentioned problem, this study aims to explore students' online reading strategies at an EMI university in Vietnam through their perceptions. To achieve the research objective, this study focuses on the following question: How do students at an EMI university use metacognitive online reading strategies?

2. Literature Review

2.1 Reading strategies

As Par (2020) stated, reading strategies are deliberate plans that readers use flexibly and adaptively to reconstruct texts. These plans involve conscious techniques aimed at improving reading comprehension on reading tasks (Kulsum et al., 2023). Reading strategies are considered important factors in comprehending text. Students can monitor their own reading and make it more efficient, and teach those strategies to others (Afflerbach et al., 2008). Moreover, readers who apply reading strategies achieve higher scores on reading comprehension tests than those who do not (Karbalaee Kamran, 2013). To summarise, learners' use of the reading strategy can predict their reading proficiency, academic achievement, and learning success (Par, 2020).

2.2 Offline reading strategies

Offline reading strategies are divided into two sub-categories: cognitive strategies and metacognitive strategies (Sheorey & Mokhtari, 2001)

2.2.1. Cognitive reading strategies

According to Sheorey and Mokhtari (2001), cognitive strategy refers to the actions or procedures used in the direct processing of text. For instance, guessing words and re-reading to better understand are two cognitive strategies. Gagne and Briggs (1979) supported this view, describing cognitive strategies as internal skills that influence the learner's intellectual processes. These processes include understanding problems, learning, remembering, and thinking.

2.2.2. Metacognitive reading strategies

Metacognitive strategies can be understood as higher-order executive skills that utilise learners' knowledge of cognitive processes. These processes help learners control their learning through activities such as planning, monitoring, and evaluating, which involve both declarative and procedural knowledge. In the same vein, Sheorey and Mokhtari (2001) argued that metacognitive strategies are intentional, carefully planned techniques that involve monitoring and management. Generally, metacognitive strategies involve two elements: the conscious awareness of strategic reading processes and the actual application of reading strategies. These strategies may distinguish successful readers from less successful readers (Sheorey & Mokhtari, 2001). In terms of categorising metacognitive reading strategies, Amer et al. (2010) classify them into three subcategories: global strategies, problem-solving strategies, and support categories. Global strategies are carefully planned to monitor students' independent reading. Problem-solving strategies are the actions or procedures readers employ when reading the text directly. Support strategies are the actions employed to aid comprehension.

2.3 Online reading strategies

Offline and online reading strategies differ due to the richness and depth of online texts (Preece, 1993). Coiro (2003) claimed that printed text is structured by the author and that readers need to follow the author's intended plot or expository structure. On the contrary, digital texts are multimodal, integrating text, static images, animations, embedded videos, and sound, and possess hyperlinks. This creates non-sequential page structures in digital texts, allowing readers to approach the content independently.

Due to the complexity of electronic texts, readers need metacognitive online reading strategies (MORS) to enhance reading comprehension. Behalova (2010) notes that online reading strategies are regarded as both learners' conscious, intentional actions and behaviours when working with electronic texts. More specifically, electronic texts are those that appear on social media platforms, in newspapers, in articles, and so on (Pookcharoen, 2009). In this study, electronic texts for learning purposes may include e-books, online articles, journals, and related materials. Generally, I suggest that MORS are intentional, carefully planned techniques used for online learning texts, including monitoring or managing the online reading process.

2.4 Perception

Perception is the process by which we absorb sensory information from our environment and use it to engage with the environment (Adiwinata, 2017). Specifically, the American Psychological Association states that perception is the process or result of becoming aware through the senses, including tasks such as recognising, observing, and discriminating. In the context of foreign language learning, learner perception refers to how students interpret themselves. In this paper, perception is defined as the process or results of IU students' self-awareness regarding the use of MORS.

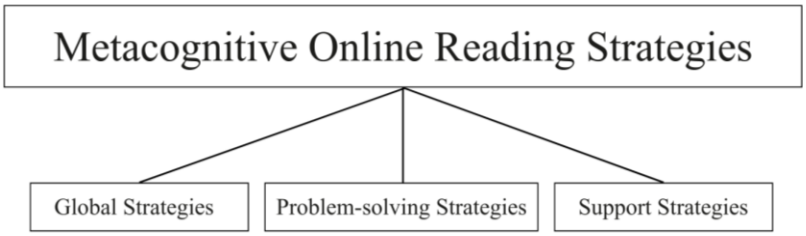
2.5 The framework

Figure 2.1 presents the framework used in this study to examine the perception of MORS. The framework is grounded in Sheorey & Mokhtari (2001) and later adapted by Amer et al. (2010). In this framework, metacognitive online reading strategies are classified into three subcategories: global, problem-solving, and support strategies.

Global strategies refer to approaches to reading texts, aiming to plan, monitor, and evaluate comprehension (Sheorey & Mokhtari, 2001). In detail, readers may apply these strategies to judge the content of the reading, attend to text features, and predict the reading topic (Anderson, 2003). Using global strategies targets readers' ability to analyse the text globally, as these strategies are significant predictors of reading performance (Rastegar et al., 2017; Yang & Sim, 2017). Problem-solving strategies are defined as strategies learners use to directly cope with challenges (Huang, 2013). As the name suggests, these strategies may provide readers with a specific plan to solve problems they may encounter while reading (Amer et al., 2010). Support strategies involve localised techniques to aid reading comprehension (Huang, 2013). In other words, these strategies involve using external resources and materials, as well as specific, useful tips (Aman & Lemu, 2023).

Figure 2.1

The framework of the study



2.5 Previous studies

Many previous studies have focused on the students' use of metacognitive online reading strategies. A mixed-methods study by Pookcharoen (2009) identified differences in MORS use between proficient and less proficient readers. In a study of Thai EFL students, the researcher found that proficient readers most frequently use problem-solving strategies. This finding aligns with Anderson's (2003), Omar's (2014), and Rianto's (2022) studies, which reported that problem-solving strategies were the most commonly used across the three subcategories, though the populations in those studies do not share similar characteristics. As they concluded, the frequent use of problem-solving strategies is shared among proficient readers and less proficient readers, ESL and EFL readers, and male and female readers. In contrast, Amer et al. (2010) investigated the use of online reading strategies among EFL students; the results revealed that global were most frequently used. Similarly, Harputra et al. (2023) found that global strategies were the most frequently used among EFL students. However, most studies similarly reported that support strategies are less frequently used than other strategies.

In addition, some studies have investigated the relation between the students' use of MORS and other aspects. Firstly, both Poole (2005) and Amer et al. (2010) examined differences in gender use of MORS. The results indicated no statistically significant difference between males and females when applying MORS. Secondly, some papers have examined the relationship between MORS use and students' reading comprehension. Most studies found a significant positive relationship between MORS use and reading comprehension (Aman & Lemu, 2023; Rastegar et al., 2017; Rianto, 2022; Zare-ee, 2007). Evidently, using MORS is necessary for students to

enhance reading comprehension when dealing with academic online texts. Thirdly, global strategies are considered powerful predictors for better comprehension (Huang et al., 2009; Yang & Sim, 2017). In particular, global strategies are identified as a significant predictor of reading performance, and excellent reading performance may imply a high level of metacognitive awareness of their use (Yang & Sim, 2017).

In Vietnam, Do and Phan (2021) conducted a study to explore the frequency with which Vietnamese undergraduates use reading strategies in academic texts. Researchers found that problem-solving strategies are most frequently used, followed by support and global strategies. As they noted, the frequency of problem-solving strategies does not differ across contexts such as Vietnam, Indonesia, and Turkey. Another finding that aligns with previous studies is the relationship between the use of global strategies and reading performance. However, no research has yet explored the students' perceived use of metacognitive online reading strategies in an EMI context. As EMI students who spend much time processing L2 texts through online materials, identifying students' perceptions may be significantly beneficial to them. Thus, this research aims to provide insights for future learners and educators to develop a comprehensive understanding of students' perceived use of metacognitive online reading strategies.

3. Methodology

3.1 Study design

The study employed a quantitative research design because it generates reliable, factual outcome data (Casebeer & Verhoef, 1997). This design helps researchers identify patterns and trends in the use of metacognitive online reading strategies (MORS) among students in an English-medium of Instruction (EMI) setting.

3.2 Participants

The study population consists of students at an English-medium instruction (EMI) university, the International University - VNU HCMC. The study sample consisted of 124 students, evenly divided between English and non-English majors. The convenience sampling method is used in this paper because it is easy to obtain as many students as possible (Triola, 2010). All participants have sufficient experience with online materials at the EMI university, which was considered appropriate for the study. To confirm voluntary participation, participants were provided with informed consent that outlined the study purpose and procedures. Demographic information for all participants is presented in Table 3.1.

Table 3.1

Demographic information of participants

Major	**Number of participants**
English majors	62
Non-English majors	62
Total	**124**

3.3 Data collection instrument

The survey was used to collect data on students' perceived use of MORS. The Online Survey of Reading Strategies (OSORS) by Amer (2010) was employed in the present study. As Amer et al. (2010) noted, this tool was adapted from the Survey of Reading Strategies (SORS) (Sheorey & Mokhtari, 2001). This adaptation was intended to differentiate offline from online reading strategies (Amer, 2010). This adaptation has made this tool appropriate for this study, which examines online rather than offline reading strategies.

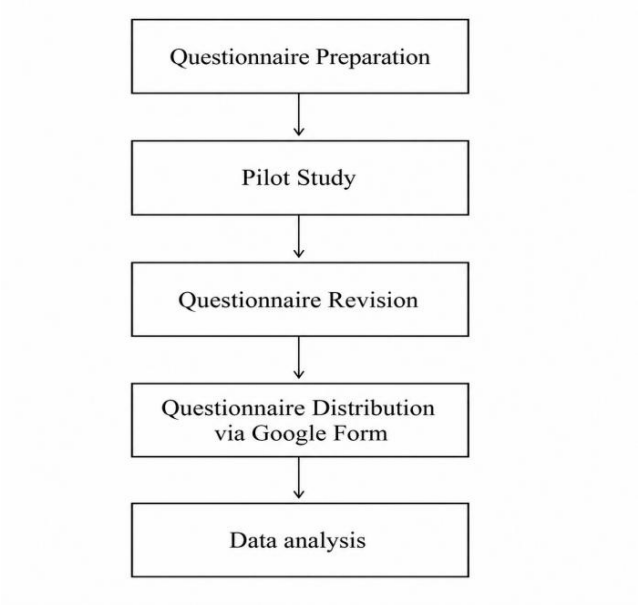
In this paper, the researcher used a Five-point Likert Scale via Google Form. Google Forms was chosen because it helped target a wide range of participants at low cost, ensured anonymity, and facilitated automatic sorting and export to SPSS at subsequent stages (Bhalerao, 2015). The survey consisted of two parts: (A) demographic information and (B) the perceived use of MORS among EMI students. The first section included four questions about confirming enrollment at EMI University, using online materials for learning, major, and gender, while the second section asked participants to rate their frequency of MORS use based on their perceptions. In the second section, there are 34 questions that indicate 34 MORS strategies, categorised into three subcategories. Items B1 - B17 surveyed students' perception of the use of global strategies, while items B18 - B25 recorded their use of problem-solving strategies, and items B26 - B34 examined their perceptions toward using support strategies. This structured distribution enabled targeted statistical analysis of each subcategory independently.

3.4 Procedure of the study

Figure 3.1 illustrates the data collection and analysis procedure. After adopting the framework and data collection instrument, the researcher conducted a pilot test to assess the clarity, validity and reliability of the instrument (Creswell & Creswell, 2018). The questionnaire was distributed to 20 volunteer participants, who had characteristics similar to those of the study sample. After the pilot test, Cronbach's alpha was calculated at .80, confirming the instrument's reliability (Taber, 2017). Additionally, two items in the global strategies have been revised based on participants' feedback to improve clarity and avoid ambiguity. In the official data collection, participants were provided with informed consent at the beginning of the survey, which outlined the nature of the study and their voluntary participation. After collecting data, Cronbach's alpha was recalculated to reconfirm reliability before data analysis. The result revealed acceptable internal consistency for the overall scale ($\alpha = .89$), confirming the survey's reliability (Taber, 2018).

Figure 3.1

Data collection and analysis procedure



3.5 Data analysis

As the study design is quantitative, this paper used the Statistical Package for the Social Sciences (SPSS) version 27 to analyse data from OSORS responses. The researcher aimed to describe the data numerically, using means (M) and standard deviations (SD) to identify trends in students' perceived use of MORS (Okamoto & Oanh, 2026). In terms of the mean, determining the central value generalises the sample mean to the population mean (Triola, 2010). By measuring the variation of values about the mean, the standard deviation indicates how much variation there is. Using these values may offer the researcher an overall view of MORS use, including the most and least frequently used strategies across the 34 OSORS items and within each sub-category. This procedure may answer the question of how students perceive their use of MORS.

To interpret the frequency of students' responses in the survey, this research uses the mean interpretation, as in Table 3.2, guided by Pimentel (2010). Regarding the interpretation of the standard deviation, Creswell's (2014) scale is used to assess variation in students' frequency of use around the mean, as shown in Table 3.3.

Table 3.2

Interpretation of the Mean Values

Means interval	Interpretation
1 - 1.50	Never
1.51 - 2.50	Rarely
2.51 - 3.50	Sometimes

3.51 - 4.50	Often
4.51 - 5	Always

Table 3.3

Interpretation of the Standard Deviation Values

Standard deviations interval	Interpretation
0.00 - 0.49	Very Low Variation (High Agreement)
0.50 – 0.99	Low to Moderate Variation
1.00 – 1.49	Moderate to High Variation
1.50 – 2.00	High Variation (Diverse Opinion)

4. Results

4.1 The perceived use of metacognitive online reading strategies

An overview of the data characteristics is presented using mean scores and SDs. Each sub-category of the metacognitive online reading strategies (MORS) is demonstrated in Table 4.1 below. As shown in the table, the problem-solving strategies have the highest mean score of 3.83, indicating that this group is the most frequently used among the 3 groups. All 3 subcategories (global, problem-solving, and support strategies) are often used by students when reading online learning materials. In other words, there is little difference in students' use of MORS at IU. Furthermore, students share a moderate perception of their use of global strategies (SD = 0.96) and problem-solving strategies (SD = 0.98). However, the SD value of 1.17 for support strategies indicates that the perceived use of support strategies falls within a different range than that of the other support strategies among IU students.

Table 4.1

Metacognitive Online Reading Strategies Means and Standard Deviations

Sub-category	Mean	SD
Problem-solving strategies	3.83	0.99
Global strategies	3.78	0.97
Support strategies	3.51	1.17

Total	**3.71**	**1.04**

4.2 The perceived use of global strategies

Regarding the students' perceived use of global strategies (B1 - B17), the overall mean is 3.78, and the SD is 0.97, indicating moderate variation in the number of students who often apply global strategies in online reading. For global strategies, item B7 has the highest mean of 3.95, indicating that students often adjust their attention to better understand online readings when encountering difficult texts. In contrast, item B3 has the lowest mean value at 3.50. Evidently, students sometimes apply the strategy of previewing online texts before reading.

Table 4.2

Students' perceived use of global strategies

Item	Statement	Mean	SD
B7	When online text becomes difficult, I pay closer attention to what I am reading.	3.95	0.94
B2	I think about what I have known to help me understand what I read online.	3.94	0.80
B9	I use context clues to help me better understand what I am reading online.	3.94	0.80
B1	I have a purpose in mind when I read online.	3.92	0.83
B12	When reading online, I check my understanding when I come across new information.	3.87	0.96
B6	When reading online, I decide what to read thoroughly and what to ignore.	3.82	0.98
B8	I use tables, figures, and pictures in the online text to increase my understanding.	3.82	1.11
B4	I think about whether the content of the online text fits my reading purpose.	3.81	0.95

B17	When reading online, I look for sites that cover both sides of an issue.	3.81	0.94
B16	I critically evaluate the online text before choosing to use its information.	3.79	0.96
B10	I use typographical features like boldface and italics to identify key information.	3.76	1.07
B13	I try to guess what the content of the online text is about when I read.	3.75	1.01
B14	I check to see if my guesses about the online text are right or wrong.	3.70	0.96
B15	I scan the online text to get a basic idea of whether it will serve my purposes before deciding to read it.	3.70	1.06
B11	I critically analyse and evaluate the information presented in the online text.	3.61	0.84
B5	I review the online text first by noting its characteristics like length and organisation.	3.60	1.04
B3	I take an overall view of the online text to see what it is about before reading it.	3.50	1.14
Total		**3.78**	**0.97**

4.3 The perceived use of problem-solving strategies

Concerning the students' perceived use of problem-solving strategies (B18-B25), the overall mean is 3.83, and the SD is 0.99, indicating a moderate gap in their perceptions of how often they use these strategies. Item B23 has the highest mean value of 4.02, indicating that students frequently use these strategies when reading online. Notably, this strategy also reaches the highest mean value among all MORS. In contrast, item B25 has the lowest mean value at 3.54. Evidently, students sometimes distinguish between facts and opinions when acquiring online texts.

Table 4.3

Students' perceived use of problem-solving strategies

Item	Statement	Mean	SD

B23	When online text becomes difficult, I re-read it to increase my understanding.	4.02	1.01
B19	I try to get back on track when I lose concentration.	3.91	0.91
B18	I read slowly and carefully to make sure I understand what I am reading online.	3.90	0.94
B22	I try to picture or visualise information to help remember what I read online.	3.90	0.92
B20	I adjust my reading speed according to what I am reading online.	3.89	0.87
B24	When I read online, I guess the meaning of unknown words or phrases.	3.83	1.06
B21	I stop from time to time and think about what I am reading online.	3.63	1.08
B25	I can distinguish between fact and opinion in online texts.	3.54	1.01
Total		**3.83**	**0.99**

4.4 The perceived use of support strategies

Regarding students' perceptions of applying support strategies (B26 - B34) for online reading, the overall mean is 3.51, and the SD is 1.17, indicating that students often use these strategies, though their opinions of their use may differ moderately. In particular, item B29 has the highest mean value at 3.93, indicating that students often use reference materials to support their online reading comprehension. In contrast, item B26 has the lowest mean at 3.05, revealing that some students may use note-taking to support their online reading comprehension. Coincidentally, this strategy is the least frequently used among all MORS by students at the EMI setting.

Table 4.4

Students' perceived use of support strategies

Item	Statement	Mean	SD
B29	I use reference materials (e.g., an online dictionary) to help me understand what I read online.	3.93	1.06
B34	When reading online, I think about information in both English and my mother tongue.	3.83	1.12

B33	When reading online, I translate from English into my native language.	3.76	1.14
B32	I ask myself questions I want to have the answers in the online text.	3.56	1.12
B31	I go back and forth in the online text to find relationships among ideas in it.	3.52	1.01
B30	I paraphrase (restate ideas in my own words) to better understand what I read online.	3.44	1.14
B27	When online text becomes difficult, I read aloud to help me understand what I read.	3.32	1.26
B28	I print out a hard copy of the online text, then underline or circle information to help me remember it.	3.17	1.40
B26	I take notes while reading online to help me understand what I read.	3.05	1.31
Total		**3.51**	**1.17**

5. Discussion

Based on the collected data, problem-solving strategies are the most frequently used among the 3 MORS sub-categories (M = 3.83). This finding aligns with Do and Phan's (2021) study of EFL students' use of metacognitive online reading strategies in a Vietnamese university context. According to them, most participants tend to use problem-solving strategies rather than other MORS sub-categories. The frequent use of these strategies may be explained by the linguistic challenges students face when acquiring academic online texts (Meniado, 2016).

Regarding the use of global strategies, this study found that item B7 had the highest mean score of 3.95. Attention adjustment during online text acquisition significantly affects the reading process, particularly reading comprehension (Yildiz & Çetinkaya, 2017). According to the automatization theory (LaBerge & Samuels, 1974), attention is an important factor for reading comprehension. Thus, students tend to use this strategy to enhance their reading comprehension. In comparison, item B3 has the lowest mean score (M = 3.50), indicating that students sometimes preview online texts before reading. As this strategy requires readers to determine the text's purpose before reading the whole text, students may comprehend it more successfully (Rukmini et al., 2014). However, if students encounter unfamiliar academic vocabulary for the first time when applying this strategy, they may find it difficult to use it, leading to discouragement and fear of its frequent use (Rukmini et al., 2014).

Regarding the use of problem-solving strategies, item B23 has the highest mean in this subcategory (M = 4.02), indicating frequent rereading to better comprehension. As Laverspool (2008) noted, rereading is recognised as one of the oldest and principal means for readers to evaluate and regulate their reading process. At various instructional levels, rereading remains one of the primary reading strategies taught to students (Laverpool, 2008). Thus, it is understandable that this strategy is most frequently used among students. In contrast, item B25 reaches the lowest mean (M = 3.54) in this group. As Graney (1990) presented, determination between opinions and facts is a critical reading problem in text evaluation. By attempting to evaluate the text, readers may recognise both explicit and implicit information (Brown & Yule, 1983). However, given the richer, more complex components of online texts, readers may find it more difficult to distinguish

between opinions and facts (Pardede, 2022). This challenge may lead to reduced confidence in applying this strategy, which explains its least frequent use among problem-solving strategies.

In terms of the use of support strategies, item B29 has the highest mean (M = 3.93), indicating frequent use of reference materials to aid reading comprehension. This result supports Hamdi (2015), who found that electronic dictionaries, a type of reference material, would be effective and motivating aids for successfully acquiring texts. Pookcharoen (2009) stated that using reference materials is one of the most frequently used strategies to address unfamiliar vocabulary or concepts. This strategy may be applied by both proficient and non-proficient readers. While non-proficient readers perceive these learning aids as the simplest way to understand the meanings of words or concepts, proficient readers use them to check meanings after using context clues and guessing. In comparison, item B26 regarding the use of notetaking to support comprehension has the lowest mean among 34 strategies (M = 3.05), which is consistent with Anderson (2003). By investigating the use of MORS among both ESL and EFL students, his study revealed that this strategy is among the least-used strategies. Judging from the EMI context of this study, students must simultaneously understand disciplinary concepts and process unfamiliar academic vocabulary to comprehend the material. Taking notes may add another layer to the complexity, in which students must identify important information, reformulate it, organise it, and write it down. This possible cause resonates with Salame et al. (2024)'s study, claiming that effective note-taking itself requires active selection and reformulation rather than simple copying. This process turned out to be more demanding for students, especially those at the beginning of their undergraduate studies. Therefore, this may explain why item B26 is the least-used strategy.

6. Conclusion

The study indicates how students at an English-medium instruction (EMI) university perceive the use of metacognitive online reading strategies (MORS), particularly the frequent use of all three sub-categories: global, problem-solving, and support strategies. Students most frequently use problem-solving strategies, aligning with other previous studies. Students frequently pay attention to the text, reread, and use reference materials to support successful reading, while applying previewing, note-taking, and determining between opinions and facts less frequently than other strategies. These findings support previous studies on students' use of metacognitive online reading strategies.

Besides, this research has some limitations. A sample size of 124 participants from an EMI university may affect the generalizability of the results. In addition, the Online Survey of Reading Strategies (OSORS) is a self-report instrument that asks participants to complete the survey based on their perceptions; the results may be subjective (Jusoh & Abdullah, 2015)

These implications suggest that lecturers, educators, and researchers may consider them for future work. Particularly, as a self-report questionnaire, OSORS cannot capture students' actual use of MORS; therefore, future research in this field should employ longitudinal designs or experimental methods to measure it more precisely. Furthermore, expanding the sample size to enhance the generalizability of the results is recommended for future studies.

Acknowledgments

First and foremost, we would like to express our sincere gratitude to International University, Vietnam National University Ho Chi Minh City, and the School of Languages for

providing the necessary support and conditions for the completion of this study. We are also deeply thankful to the students who generously participated in the research. Their honest responses and valuable time made an important contribution to the success of this study. Our appreciation is further extended to the teachers who kindly allowed us to collect data in their classrooms. Finally, we would like to thank the conference chairs and reviewers for their insightful and constructive comments on earlier versions of this article.

References

Adiwinata, A. (2017). The students' perception towards the effectiveness of public speaking subject to support their speaking skill. Makassar Muhammadiyah University

Afflerbach, P., Pearson, P. D. & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*(5), 364–373. https://doi.org/10.1598/RT.61.5.1

Ally, M. (2013). Mobile learning: From research to practice to impact education. *Learning and Teaching in Higher Education: Gulf Perspectives, 10*(2), 3–12. https://doi.org/10.18538/lthe.v10.n2.140

Aman, M. (2024). *Exploring the relation between reading strategy use and reading comprehension among first-year Wollo University students* (Doctoral dissertation, Addis Ababa University). Addis Ababa University.

Aman, M. & Lemu, G. (2023). Exploring reading strategy use of Wollo University first-year social science students. *Abyssinia Journal of Business and Social Sciences, 8*(2), 73-83. https://doi.org/10.20372/ajbs.2023.8.2.887

Amer, A. A., Barwani, T. A. A., & Ibrahim, M. M. (2010). Student teachers' perceived use of online reading strategies. *International Journal of Education and Development using Information and Communication Technology, 6*(4), 102-113. https://www.learntechlib.org/p/42266/

Anderson, N. J. (2003). Scrolling, clicking, and reading English: Online reading strategies in a second/foreign language. *The Reading Matrix, 3*(3), 1–33. http://www.readingmatrix.com/articles/anderson/article.pdf

AYDINBEK, C. (2021). The effect of instruction in reading strategies on the reading achievement of learners of French. *Eurasian Journal of Educational Research, 21*(91), 321–338. https://doi.org/10.14689/ejer.2021.91.15

Behalova, A. (2010). Exploring online reading strategies of American undergraduate students (Doctoral dissertation, Oklahoma State University).

Bhalerao, A. K. (2015). Application and performance of Google Forms for online data collection and analysis: A case of North Eastern Region of India. *Indian Journal of Extension Education, 51*(3 & 4), 49–53. https://indianjournals.com/article/ijee3-51-3and4-010

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge University Press.

Carrell, P. L., Pharis, B. G. & Liberto, J. C. (1989). Metacognitive strategy training for ESL reading. *TESOL Quarterly, 23*(4), 647–678. https://doi.org/10.2307/3587536

Casebeer, A. L. & Verhoef, M. J. (1997). Combining qualitative and quantitative research methods: Considering the possibilities for enhancing the study of chronic diseases. *Chronic Diseases in Canada, 18*(3), 130–135. https://www.floppybunny.org/robin/web/virtualclassroom/chap5/s5/articles_qual_quant_approaches/combining_qual_quant.pdf

Coiro, J. (2003). Reading comprehension on the internet: expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher, 56*(5), 458-464. https://www.jstor.org/stable/20205224

Creswell, J. W., & Creswell, J. D. (2014). *Research design*. SAGE Publications Ltd.

Creswell, J. W. & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Do, H. M. & Phan, H. L. T. (2021). Metacognitive awareness of reading strategies on second language Vietnamese undergraduates. *Arab World English Journal, 12*(1), 90–112. https://doi.org/10.24093/awej/vol12no1.7

Gagne, R. M. & Briggs, L. J. (1979). *Principles of Instructional Design*. Holt, Rinehart, and Winston.

Graney, J. M. (1990). Determination of fact and opinion: a critical reading problem. *Journal of Psycholinguistic Research, 19*(3), 147-166. https://doi.org/10.1007/bf01077414

Hamdi, C. (2015). The Effects of Electronic Dictionary Use on Reading Comprehension and Vocabulary Retention of EFL Students. *Arab World English Journal*, 180-191. https://www.awej.org/images/conferences/BejaiaProceedings2015/16.pdf

Harputra, Y., Rahmadani, Y. R. & Siregar, M. N. H. (2023). Identifying metacognitive online reading strategies of students Universitas Graha Nusantara. *ETANIC Journal of English Language Teaching and Applied Linguistics*, 1(1), 20–30. https://jurnal.radisi.or.id/index.php/JournalETANIC/article/view/257

Huang, H.-C. (2013). Online reading strategies at work: What teachers think and what students do. *ReCALL, 25*(3), 340–358. https://doi.org/10.1017/S0958344013000153

Huang, H. C., Chern, C. L. & Lin, C. C. (2009). EFL learners' use of online reading strategies and comprehension of texts: An exploratory study. *Computers & Education, 52*(1), 13–26. https://doi.org/10.1016/j.compedu.2008.06.003

Jusoh, Z. & Abdullah, L. (2015). Online survey of reading strategies (OSORS): Students' online reading in academic context. *Malaysian Journal of Distance Education, 17*(2), 67–81. https://doi.org/10.21315/mjde2015.17.2.5

Karbalaee Kamran, S. (2013). Does reading strategy use predict and correlate with reading achievement of EFL learners? *International Journal of Research Studies in Language Learning, 2*(2), 29–38. https://doi.org/10.5861/ijrsll.2012.118

Kulsum, E. M., Fauzan, E., & Mutiarawati, R. (2023). The use of reading strategies in EFL learning. *Gunung Djati Conference Series, 28*, 100-108. https://conferences.uinsgd.ac.id/

Kung, F.-W. (2017). Teaching second language reading comprehension: the effects of classroom materials and reading strategy use. *Innovation in Language Learning and Teaching, 13*(1), 93-104. https://doi.org/10.1080/17501229.2017.1364252

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Laverpool, A. (2008). The efficacy of rereading as a metacognitive tool for reading comprehension monitoring. *Journal of College Reading and Learning, 38*(2), 31–48. https://doi.org/10.1080/10790195.2008.10850307

Meniado, J. C. (2016). Metacognitive reading strategies, motivation, and reading comprehension performance of Saudi EFL students. *English Language Teaching, 9*(3), 117-129. https://doi.org/10.5539/elt.v9n3p117

Okamoto, S. I., & Oanh, N. T. K. (2025). Statistics in environmental monitoring and assessment. Chapman and Hall/CRC. https://doi.org/10.1201/9781003597216

Omar, N. A. (2014). Online metacognitive reading strategies use by postgraduate Libyan EFL students. *International Journal of Humanities and Social Sciences, 8*(7), 2281–2284. https://doi.org/10.5281/zenodo.1093956

Par, L. (2020). The relationship between reading strategies and reading achievement of the EFL students. *International Journal of Instruction, 13*(2), 223–238. https://doi.org/10.29333/iji.2020.13216a

Pardede, P. (2022). Online reading strategies in EFL: A review. *JET (Journal of English Teaching), 8*(2), 329–339. https://doi.org/10.33541/jet.v8i2.4130

Parker, K., & Lenhart, A. (2011). *The digital revolution and higher education: College presidents, public differ on value of online learning*. Pew Research Center. https://www.pewresearch.org/social-trends/2011/08/28/the-digital-revolution-and-higher-education/

Perception. (n.d.). In APA dictionary of psychology. American Psychological Association Dictionary of Psychology.

Phuong, Y. H. & Nguyen, T. T. (2019). Students' perceptions towards the benefits and drawbacks of EMI classes. *English Language Teaching, 12*(5), 88–100. https://doi.org/10.5539/elt.v12n5p88

Pimentel, J. L. (2010). A note on the usage of Likert scaling for research data analysis. *USM R&D Journal, 18*(2), 109–112.

Poole, A. (2005). Gender differences in reading strategy use among ESL college students. *Journal of College Reading and Learning, 36*(1), 7–20. https://doi.org/10.1080/10790195.2005.10850177

Pookcharoen, S. (2009). *Metacognitive online reading strategies among thai EFL university students*. ProQuest LLC. https://eric.ed.gov/?id=ED515675

Preece, J. (1993). Hypermedia, multimedia and human factors. In C. Latchem, J. Williamson, & L. Henderson-Lancett (Eds.), *Interactive multimedia: Practice and promise* (pp. 135–150). Kogan Page.

Putera, A. A. & Anggrainy, F. P. N. (2024). Online reading strategies to boost the students' motivation. *Journal of English Language Teaching and Learning (JETLE), 5*(2), 173–184. https://doi.org/10.18860/jetle.v5i2.26522

Raja, R. & Nagasubramani, P. C. (2018). Impact of modern technology in education. *Journal of Applied and Advanced Research, 3*(Suppl. 1), S33–S35. https://doi.org/10.21839/jaar.2018.v3is1.165

Rastegar, M., Kermani, E. M., & Khabir, M. (2017). The relationship between metacognitive reading strategies use and reading comprehension achievement of EFL learners. *Open Journal of Modern Linguistics, 07*(02), 65-74. https://doi.org/10.4236/ojml.2017.72006

RIANTO, A. (2022). EXPLORING CORRELATION BETWEEN METACOGNITIVE ONLINE READING STRATEGY USE AND ONLINE READING COMPREHENSION OF EFL STUDENTS. *Turkish Online Journal of Distance Education, 23*(2), 223-235. https://doi.org/10.17718/tojde.1096448

Rukmini, N., Ohoiwutun, J. E. & Muhsin. (2014). Improving reading comprehension of the grade eight students by applying previewing and predicting technique. *e-Journal of English Language Teaching Society (ELTS), 2*(4), 1–11.

Salame, I. I., Tuba, M., & Nujhat, M. (2024). Note-taking and its impact on learning, academic performance, and memory. *International Journal of Instruction, 17*(3), 599-616. https://doi.org/10.29333/iji.2024.17333a

Sheorey, R. & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System, 29*(4), 431–449. https://doi.org/10.1016/S0346-251X(01)00039-2

Taber, K. S. (2017). The use of cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Triola, M. F. (2010). *Elementary statistics* (11th ed.). Pearson Education.

Wang, Y. H. (2016). Reading strategy use and comprehension performance of more successful and less successful readers: A think-aloud study. *Educational Sciences: Theory & Practice, 16*(5), 1789–1813. https://doi.org/10.12738/estp.2016.5.0116

Yang, L., & Sim, T. S. (2017). Metacognitive awareness of reading strategies among EFL high school students in China. *AJELP: The Asian Journal of English Language and Pedagogy, 5*, 34-45. https://doi.org/10.37134/ajelp.vol5.4.2017

Yildiz, M. & Çetinkaya, E. (2017). The relationship between good readers' attention, reading fluency and reading comprehension. *Universal Journal of Educational Research, 5*(3), 366–371. https://doi.org/10.13189/ujer.2017.050309

Zare-ee, A. (2007). The relationship between cognitive and meta-cognitive strategy use and EFL reading achievement. *Journal of Applied Psychology, 2*(5), 105–119.

Bionote

Thinh Tien Vu is head of the Testing board and vice dean of the School of Languages, International University – VNU HCM. He received his M.A. in TESOL from the University of Canberra, Australia, and is currently a PhD candidate in English Language Studies and Teaching at NIDA, Thailand. He has 25 years of experience in teaching academic English, ESP, test preparation, and TESOL courses, and more than 12 years of experience in the roles of course/program designer and developer, academic trainer, and manager. His current research interests include EMI, Applied methodologies, Autonomy development, Assessment and Testing, Education management, and TELL.

Thi Thuy Dung Dang is an undergraduate student majoring in English Linguistics and Literature at International University – VNU HCM. She is currently teaching English to young learners while actively pursuing her academic research interests. Her academic interests include English online learning, technology-enhanced language learning (TELL), and teaching reading.

#126 LEARNERS' PERCEPTIONS OF TEACHER ORAL FEEDBACK IN VSTEP SPEAKING PREPARATION: IMPLICATIONS FOR ESL-ORIENTED INSTRUCTION IN VIETNAM

Doan Ngoc Sang[1], Pham Huu Duc [2]

Abstract

Teacher oral feedback plays a crucial role in supporting learners' speaking preparation, particularly in high-stakes test contexts. This study investigates learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes at Lac Hong University, Vietnam. Data were collected from 64 students through a structured questionnaire measuring four dimensions of feedback: clarity, timing, focus, and affective tone. The findings indicate that learners generally hold positive perceptions of teacher oral feedback (Mean = 3.90). Among the dimensions, feedback clarity and focus were rated highest, suggesting that learners value explicit and task-related guidance. In contrast, affective tone received lower ratings, with evaluative feedback negatively associated with learners' reported confidence. Correlation analysis further revealed that feedback clarity was positively associated with perceived speaking development, whereas supportive feedback was positively associated with learners' perceived confidence and willingness to communicate. These findings highlight the importance of balancing instructional clarity and emotional support in speaking feedback practices. Importantly, the observed relationships are correlational and should not be interpreted as evidence of causal effects. The study suggests that effective oral feedback in VSTEP speaking classrooms may support learners' perceived speaking development and communicative confidence, providing pedagogical insights for the transition toward ESL-oriented instruction in Vietnam. The study contributes to ongoing discussions on language teaching in Vietnam, where English is increasingly positioned as a second language.

***Keywords*:** vstep; oral feedback; learner perceptions; speaking skills; esl-oriented instruction

1. Introduction

The increasing institutionalization of English proficiency requirements in higher education has positioned standardized assessments as central drivers of language teaching and learning practices, as widely noted in research on assessment driven education systems (Fulcher, 2024). In Vietnam, the Vietnamese Standardized Test of English Proficiency (VSTEP), aligned with the CEFR, has not only functioned as a graduation benchmark but has also reshaped instructional priorities, particularly by intensifying the focus on test oriented speaking preparation, a trend commonly observed in high-stakes testing contexts where assessment exerts strong washback effects on teaching and learning. Within such high-stakes contexts, speaking performance is often

[1] *Corresponding author, Lecturer, Lac Hong Foreign Languages and Informatics Center, Lac Hong University, Dong Nai, Vietnam; Email: sangdn@lhu.edu.vn*
[2] *Associate Professor, International University, Vietnam National University–Ho Chi Minh City, Vietnam*

constrained by real-time processing demands while simultaneously being influenced by affective variables such as anxiety and fear of negative evaluation, which have been shown to significantly mediate learners' oral production MacIntyre and Gardner (1989); Dewaele & Li (2020).

Within this pedagogical landscape, teacher oral feedback remains a critical instructional mechanism, yet its effectiveness is increasingly understood as contingent upon learner interpretation rather than mere teacher provision. While earlier models conceptualized feedback as a unidirectional transmission of corrective information, more recent perspectives emphasize that learners actively construct meaning from feedback through cognitive and affective engagement, thereby determining its ultimate impact on learning outcomes (Winstone & Carless, 2021; Tai et al., 2018). This shift toward a learner-centered understanding of feedback aligns with findings that feedback clarity, relevance, and timing function as key determinants of perceived usefulness, whereas affective tone can either facilitate or hinder learner uptake depending on how feedback is emotionally received (Mäkipää, 2024; Han & Hyland, 2015).

In speaking instruction, the role of oral feedback becomes even more complex due to its immediate and interactional nature, which requires teachers to balance corrective intent with communicative flow. Empirical evidence suggests that feedback delivered during or after speaking tasks can influence learners' confidence and willingness to communicate, with supportive feedback promoting engagement while overly evaluative feedback may increase anxiety and reduce participation (Sato & Loewen, 2019; Zhang, 2018). At the same time, learners' prior experiences, beliefs, and contextual expectations shape how feedback is interpreted, reinforcing the argument that "what learners make of feedback" is as important as "what teachers provide" (Carless & Boud, 2021).

Parallel to these developments, the integration of digital technologies and artificial intelligence into language education has introduced new forms of feedback that are increasingly immediate, individualized, and data driven. Such technological affordances, while expanding opportunities for speaking practice and feedback, also raise important pedagogical questions regarding the alignment between automated feedback and human-mediated feedback processes, particularly in contexts where communicative competence remains the primary instructional goal (Godwin-Jones, 2022; Stickler & Hampel, 2022).

Despite the expanding body of research on feedback and learner engagement, previous studies have largely examined feedback in general second and foreign language learning contexts. Comparatively little attention has been paid to how learners interpret teacher oral feedback in high-stakes VSTEP speaking preparation in Vietnam, where feedback serves not only as instructional guidance but also as an indicator of test readiness. This context is particularly important because learners' interpretations of feedback may shape both their cognitive engagement and emotional responses during speaking preparation. Moreover, limited research has explored how learners' perceptions of different feedback dimensions relate to their perceived speaking development and communicative confidence in VSTEP classrooms. Addressing this gap is important because English is increasingly conceptualized not merely as a foreign language but as a second language within Vietnamese educational discourse. Consequently, there is a growing need to understand how feedback practices in test preparation can also support broader communicative learning goals aligned with ESL-oriented instruction, in which language use extends beyond examination performance to authentic communication Doan et al. (2018).

Against this backdrop, the present study investigates learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes at Lac Hong University. Specifically, it examines how learners perceive four key feedback dimensions, namely clarity, timing, focus, and affective tone, and explores how these perceptions are associated with learners' perceived speaking development and communicative confidence. By focusing on learners' interpretations of teacher oral feedback in a high-stakes VSTEP context, the study extends previous feedback research beyond general language learning settings. Furthermore, it contributes to the growing discussion on how feedback practices in test-oriented classrooms may provide pedagogical support for the broader transition toward ESL-oriented communicative instruction in Vietnam.

Therefore, this study focuses on investigating learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes, with particular attention to key feedback dimensions, including clarity, timing, focus, and affective tone. Given the limited empirical evidence in high-stakes test preparation contexts in Vietnam, examining how learners interpret and respond to oral feedback is essential for understanding its pedagogical effectiveness. To bridge these gaps, this study attempts to answer the following research questions:

RQ1: What are learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes in terms of clarity, timing, focus, and affective tone?

RQ2: How do these perceptions influence learners' speaking development and communicative confidence in a high-stakes test preparation context?

2. Literature Review

2.1. Feedback Theory

Feedback has been widely acknowledged as a central mechanism in second language learning, functioning not only as a tool for correcting linguistic errors but also as a means of guiding learners' cognitive and metacognitive development (Hattie & Timperley, 2007; Ellis, 2020). While earlier perspectives emphasized feedback as information transmission, more recent frameworks conceptualize feedback as an interactive and learner-mediated process in which meaning is actively constructed rather than passively received (Carless & Boud, 2018; Winstone & Carless, 2019). Within this reconceptualization, feedback effectiveness depends on how learners interpret and utilize the information, as feedback only becomes meaningful when it is actively taken up and applied in subsequent learning processes (Tai et al., 2018; Zhang & Hyland, 2018).

Extending this view, Tai et al. (2018) highlight that feedback operates across multiple levels, including task, process, and self-regulation, each contributing differently to learning outcomes depending on how learners engage with them. This multidimensional nature of feedback is further supported by empirical evidence showing that clarity and alignment with learning goals significantly influence learners' ability to use feedback productively, while unclear or unfocused feedback often leads to disengagement. In this regard, Mäkipää (2024) demonstrates that learners who perceive feedback as clear and actionable are more likely to maintain engagement and regulate their learning effectively.

At the same time, recent studies increasingly emphasize the affective dimension of feedback, suggesting that emotional responses play a crucial role in shaping learners' engagement. While feedback has traditionally been associated with cognitive gains, it is now recognized that factors such as confidence, anxiety, and motivation mediate whether feedback leads to

improvement. For instance, Zhang (2018) shows that feedback delivered within a supportive pedagogical environment enhances learner engagement, whereas overly evaluative feedback may undermine motivation and reduce participation. Similarly, Han and Hyland (2015) argue that the emotional tone of feedback significantly influences how learners process and respond to corrective input.

Taken together, these theoretical perspectives provide the foundation for the four feedback dimensions examined in the present study. Specifically, feedback clarity refers to the extent to which learners perceive feedback as understandable and actionable (Hattie & Timperley, 2007; Mäkipää, 2024). Feedback timing concerns when feedback is delivered in relation to speaking tasks and how its immediacy influences learner engagement (Sato & Loewen, 2019). Feedback focus reflects the extent to which feedback targets specific aspects of task performance and learning goals (Tai et al., 2022). Finally, affective tone captures the emotional quality of teacher feedback and its potential role in shaping learners' confidence and willingness to engage in speaking activities (Han & Hyland, 2015; Zhang, 2018). These four dimensions therefore represent complementary cognitive and affective aspects of teacher oral feedback that are particularly relevant to speaking instruction.

2.2. Learner Perception of Feedback

A growing body of research has shifted the focus of feedback studies toward learners' perceptions, recognizing that feedback does not inherently lead to learning unless it is meaningfully interpreted by learners. This learner centered perspective posits that feedback effectiveness is mediated by learners' beliefs, prior experiences, and contextual expectations, rather than being determined solely by teacher practices. As Carless and Boud (2021) emphasize, feedback should be understood as a process of sense-making in which learners actively engage with information to regulate their learning.

Recent empirical studies further demonstrate that learners' perceptions are shaped by both cognitive and affective factors. For example, Zhang and Hyland (2018) found that learners are more likely to engage with feedback when it is perceived as relevant, specific, and actionable, whereas vague or overly critical feedback can hinder engagement. Complementing this view, research on feedback literacy suggests that learners need specific competencies to interpret and use feedback effectively, including the ability to manage emotional responses and translate feedback into action (Tai et al., 2018).

Moreover, emerging studies highlight the importance of affective responses in shaping feedback perception. Learners who perceive feedback as supportive tend to demonstrate greater willingness to communicate and participate in learning activities, while negative perceptions may lead to increased anxiety and avoidance behaviors. For instance, recent findings indicate that emotional engagement plays a decisive role in determining whether learners act upon feedback, reinforcing the idea that feedback effectiveness is inseparable from learners' affective experiences. Accordingly, the present study adopts these four dimensions as the conceptual basis of the questionnaire because they represent key characteristics through which learners interpret teacher oral feedback. Examining learners' perceptions across these dimensions provides a learner-centered perspective on feedback effectiveness, complementing previous studies that have primarily focused on teachers' feedback practices.

2.3. Oral Feedback in Speaking and VSTEP Contexts

Oral feedback occupies a distinctive position in speaking instruction due to its immediacy and interactional nature, requiring learners to process feedback while simultaneously engaging in communication. Unlike written feedback, which allows for delayed reflection, oral feedback is embedded within real-time interaction, making it both a powerful and potentially disruptive instructional tool. Research indicates that different forms of oral feedback, including explicit correction and recasts, play varying roles in supporting language development, with explicit feedback often facilitating greater noticing of linguistic forms.

The effectiveness of oral feedback in speaking contexts depends on how teachers balance corrective intent with communicative flow. While feedback delivered during interaction can promote immediate awareness, it may also interrupt fluency if not managed carefully. Conversely, delayed feedback may support reflection but reduce immediacy of learning. Sato and Loewen (2019) argue that optimal feedback practices require a dynamic balance between these competing demands, highlighting the importance of contextual sensitivity in feedback delivery.

In high-stakes test preparation contexts such as VSTEP, oral feedback assumes an additional evaluative dimension, as it is often interpreted by learners as an indicator of test readiness. This dual role both instructional and evaluative can shape learners' perceptions in complex ways, influencing not only their cognitive engagement but also their emotional responses. Despite the widespread implementation of VSTEP in Vietnamese higher education, research examining learners' perceptions of oral feedback in this context remains limited, indicating a clear gap in the literature.

Recent developments in educational technology have introduced AI-mediated feedback as an additional source of speaking support. Although previous studies suggest that AI can provide immediate and individualized feedback (Xu et al., 2024), the present study focuses specifically on teacher oral feedback, given its interactive, contextual, and relational nature in VSTEP speaking classrooms. Therefore, AI-mediated feedback is acknowledged as a complementary development rather than the primary focus of this research.

Based on the literature reviewed above, the present study operationalizes teacher oral feedback through four dimensions: clarity, timing, focus, and affective tone. These dimensions were selected because they consistently emerge in contemporary feedback research as key factors influencing how learners interpret, engage with, and respond to teacher feedback in speaking classrooms. They also provide the conceptual framework for the questionnaire used to investigate learners' perceptions in the present study.

3. Methodology

3.1. Research Design

This study employed a quantitative survey design to investigate learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes. A quantitative approach was considered appropriate because the primary objectives were to describe learners' perceptions across four predefined feedback dimensions and to examine the statistical relationships between these dimensions and learners' perceived speaking development and perceived confidence. Since the study focused on identifying general patterns of learners' responses rather than exploring

individual experiences in depth, a structured questionnaire was considered the most suitable instrument.

Although a mixed-methods design incorporating interviews might provide richer insights into how learners interpret teacher feedback, the present study adopted a quantitative approach to establish an initial empirical understanding of learners' perceptions within a high-stakes VSTEP context. Future studies may complement these findings through qualitative or mixed-methods research to provide a deeper explanation of learners' feedback experiences.

3.2. Participants

This study was conducted at the Lac Hong Foreign Languages and Informatics Center, Lac Hong University, Dong Nai Province, Vietnam. The participants were 64 undergraduate students enrolled in VSTEP B1-B2 speaking preparation courses. Participants were recruited from two intact VSTEP speaking preparation classes (K33 and K34) using convenience sampling. Participation was voluntary, and all respondents completed the questionnaire anonymously.

Table 1

Summarizes the demographic characteristics of the participants.

Characteristic	**Category**	**n**	**%**
Gender	Male	32	50.0
	Female	32	50.0
VSTEP Preparation Course	K33	29	45.3
	K34	35	54.7

Basic demographic information, including gender and course enrollment, was collected to describe the sample. Additional background variables, such as academic year, previous VSTEP experience, and preparation duration, were not collected and are therefore acknowledged as limitations of the study. As the sample was drawn from a single language center within one university, the findings should be interpreted with caution and cannot be generalized to all VSTEP learners in Vietnam.

3.3. Instrument

Data were collected using a structured questionnaire specifically designed to measure learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes. The questionnaire consisted of 20 Likert-scale items (five items for each feedback dimension) measured on a five-point scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

The four dimensions were developed based on established feedback literature.

- Feedback clarity focused on the extent to which feedback was understandable, specific, and actionable, drawing on feedback usability (Hattie & Timperley, 2007; Mäkipää, 2024).

- Feedback timing examined learners' perceptions of when feedback was provided during or after speaking activities, based on interactional feedback research (Sato & Loewen, 2019).

- Feedback focus measured whether teacher feedback addressed specific aspects of speaking performance and learning goals, reflecting task alignment (Tai et al., 2018).

- Affective tone investigated learners' perceptions of teachers' emotional support, encouragement, and sensitivity when providing feedback, drawing upon research on learner engagement and feedback emotion (Han & Hyland, 2015; Zhang, 2018).

Sample questionnaire items included:

- "My teacher's feedback is easy to understand."
- "Feedback is provided at an appropriate time."
- "Teacher feedback focuses on important speaking problems."
- "My teacher gives feedback in an encouraging manner."

The questionnaire was adapted from contemporary feedback perception literature and further contextualized for VSTEP speaking preparation through a review of relevant theoretical frameworks and recent empirical studies.

3.4. Instrument Reliability and Validity

Prior to the main survey, the questionnaire was reviewed by experienced VSTEP instructors and researchers in language education to ensure content validity and contextual appropriateness. Minor revisions were made to improve wording clarity and item specificity. Internal consistency was evaluated using Cronbach's alpha. All four constructs exceeded the commonly accepted threshold of .70, indicating satisfactory reliability.

Table 2

Participant Characteristics (N = 64)

Construct	**Cronbach's α**
Clarity	.88
Timing	.82
Focus	.86
Affective Tone	.84

Overall questionnaire α = .90. These results indicate satisfactory internal consistency for all constructs included in the questionnaire.

3.5. Data Analysis

The collected data were analyzed using SPSS. Descriptive statistics (minimum, maximum, mean, and standard deviation) were calculated to summarize learners' perceptions across the four feedback dimensions. Pearson correlation analysis was then conducted to examine the relationships between these feedback dimensions and learners' perceived speaking development as well as perceived confidence. Because both outcome variables were measured through self-report questionnaire items, the findings indicate statistical associations rather than causal relationships.

This study was conducted at Lac Hong University (LHU) in Dong Nai Province, Vietnam. It focuses on learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes, particularly in relation to key feedback dimensions, including clarity, timing, focus, and

affective tone. The study investigated a group of 64 undergraduate students who were preparing for the VSTEP B1-B2 speaking test at the Lac Hong Foreign Languages and Informatics Center.

A quantitative research approach was employed to collect and analyze data from the participants. Specifically, a structured questionnaire was used as the primary research instrument, as it is widely recognized as an effective tool for gathering large-scale perception data and examining patterns of learner responses in educational contexts (Winstone & Carless, 2019). The design of the questionnaire was informed by established feedback frameworks, particularly those emphasizing the role of feedback clarity, timing, and learner engagement in shaping learning outcomes (Tai et al., 2018). In addition, the inclusion of affective dimensions in the instrument reflects recent findings that emotional responses to feedback significantly influence learners' willingness to engage with speaking tasks (Han & Hyland, 2015).

Participants were asked to evaluate their perceptions of teacher oral feedback in speaking classes using a Likert-scale questionnaire. The instrument consisted of multiple items measuring different aspects of feedback, allowing for a comprehensive understanding of how learners experience and interpret feedback in a high-stakes test preparation context. This approach aligns with previous research suggesting that learners' perceptions play a mediating role in determining the effectiveness of feedback practices (Carless & Boud, 2021). The questionnaire was administered to all 64 participants, and all responses were collected for analysis.

To ensure the reliability of the instrument, internal consistency was examined using Cronbach's alpha, with all constructs meeting acceptable reliability thresholds. The collected data were then analyzed using the Statistical Package for the Social Sciences (SPSS). Descriptive statistics, including minimum, maximum, mean, and standard deviation, were calculated to identify general trends in learners' perceptions. In addition, correlation analysis was conducted to examine the relationships between different feedback dimensions and learners' perceived speaking development and confidence.

The findings derived from the quantitative data provide insights into how learners perceive teacher oral feedback and how these perceptions may influence their engagement and performance in VSTEP speaking preparation.

4. Results

4.1. Descriptive Statistics of Learners' Perceptions

To address the first research question, descriptive statistics were computed based on students' responses to the questionnaire, focusing on four key dimensions of teacher oral feedback: clarity, timing, focus, and affective tone. Overall, the findings suggest that learners tend to hold positive perceptions of teacher feedback in VSTEP speaking classes, with an average mean score of 3.90.

Table 3

Descriptive Statistics of Feedback Dimensions

Dimension	N	Min	Max	Mean	Std. Deviation
Clarity	64	2.80	5.00	4.12	0.56
Focus	64	2.75	5.00	4.05	0.60
Timing	64	2.50	5.00	3.85	0.64
Affective Tone	64	2.30	5.00	3.60	0.72
Overall	64	2.70	5.00	3.90	0.58

Note. N = 64; values are based on a 5-point Likert scale.

As shown in Table 3, feedback clarity and feedback focus received the highest mean scores, indicating that learners particularly valued oral feedback that was clear, explicit, task-related, and easy to apply in subsequent speaking activities. By contrast, affective tone received the lowest mean score, suggesting greater variation in how learners perceived the emotional quality of teacher feedback. Overall, these findings indicate that learners generally held positive perceptions of teacher oral feedback, especially when feedback was understandable and directly related to speaking tasks.

By contrast, feedback timing was rated at a moderate level (Mean = 3.85), reflecting some differences in learners' preferences regarding when feedback should be provided during or after speaking tasks. The lowest mean score was observed for affective tone (Mean = 3.60), which may indicate that students have more varied experiences in terms of how feedback is delivered emotionally. This is further reflected in the relatively higher standard deviation (SD = 0.72), suggesting greater variability in learners' responses to the emotional dimension of feedback.

4.2. Correlation Analysis

To address the second research question, Pearson correlation analysis was conducted to examine the relationships between different dimensions of teacher oral feedback and learners' perceived speaking development as well as their perceived confidence. The results reveal several statistically significant associations among the variables.

Table 4

Correlation between Feedback Dimensions and Learning Outcomes

Variables	Perceived Speaking Development	Perceived Confidence
Clarity	.62**	.58**
Focus	.59**	.55**
Timing	.41*	.38*
Affective Tone	.35*	.52**

Note. The values of p indicate the level of statistical significance of the correlations. A result is considered statistically significant at p < .05, meaning there is less than a 5% probability that the relationship occurred by

*chance. A stronger level of significance is indicated by p < .01 (marked by **), suggesting a higher level of confidence in the reliability of the results.*

As shown in Table 4, feedback clarity demonstrated a strong positive correlation with learners' perceived speaking development ($r = .62$, $p < .01$) and perceived confidence ($r = .58$, $p < .01$). Similarly, feedback focus was positively correlated with both perceived speaking development ($r = .59$, $p < .01$) and perceived confidence ($r = .55$, $p < .01$). These findings indicate that learners who perceived teacher feedback as clearer and more task-focused also tended to report higher levels of perceived speaking development and confidence.

By comparison, feedback timing showed a moderate positive correlation with perceived speaking development ($r = .41$, $p < .05$), whereas affective tone demonstrated a stronger correlation with perceived confidence ($r = .52$, $p < .01$) than with perceived speaking development ($r = .35$, $p < .05$). These associations suggest that different dimensions of teacher oral feedback may relate to different aspects of learners' perceptions. However, because the study employed a correlational design based on self-reported questionnaire data, these relationships should not be interpreted as evidence of causal effects.

5. Discussion

The findings of this study provide a more nuanced understanding of learners' perceptions of teacher oral feedback in VSTEP speaking preparation, directly addressing the two research questions.

With regard to the first research question, the results indicate that learners generally hold positive perceptions of teacher oral feedback, particularly in terms of clarity and focus. These dimensions were not only rated highest but also consistently associated with learners' sense of usefulness and relevance. This suggests that learners value feedback that is explicit, structured, and closely aligned with task requirements, as it enables them to better understand performance expectations and identify areas for improvement. This finding supports Winstone and Carless (2019), who emphasize that feedback only becomes effective when learners can meaningfully interpret and act upon it. Similarly, Tai et al. (2018) highlight the importance of feedback that enhances learners' evaluative judgment and task understanding. This pattern may be partly explained by the nature of VSTEP as a high-stakes speaking examination. Because learners are preparing for an assessment in which speaking performance is evaluated against clearly defined criteria, they are likely to place greater value on feedback that is explicit, task-focused, and immediately applicable. Clear guidance may help learners better understand performance expectations, identify specific weaknesses, and prioritize areas for improvement during test preparation. In this context, clarity and focus appear to be particularly meaningful because they reduce uncertainty and provide learners with practical directions for subsequent speaking practice.

However, learners' perceptions were less consistent with regard to timing and affective tone, indicating that feedback is experienced not only as a cognitive resource but also as a socially and emotionally mediated process. In particular, the relatively lower rating of affective tone suggests that learners may be sensitive to how feedback is delivered, especially in high-pressure speaking contexts. This aligns with Han and Hyland (2015), who argue that feedback is inherently relational, and that its effectiveness depends on how it is emotionally received by learners.

In relation to the second research question, the findings reveal that different feedback dimensions play distinct roles in shaping learners' perceived speaking development and communicative confidence. Specifically, clarity and focus demonstrate strong relationships with speaking development, suggesting that cognitively oriented feedback is more influential in helping learners improve their performance. By contrast, affective tone shows a stronger association with confidence, highlighting the role of emotional support in fostering learners' willingness to communicate. This indicates that while cognitive clarity facilitates learning, emotional factors are equally important in sustaining engagement, particularly in speaking activities that require real-time performance.

The role of timing appears to be more moderate, suggesting that although it contributes to learning, it is less decisive compared to clarity and focus. This finding reflects the dynamic nature of oral interaction, where feedback must be carefully integrated into communicative flow. As Sato and Loewen (2019) suggest, the effectiveness of feedback depends not only on its content but also on its delivery within interactional contexts.

Taken together, these findings highlight the dual nature of feedback as both a cognitive and affective process, particularly in high-stakes test preparation environments. In such contexts, feedback often serves both instructional and evaluative functions, which may influence how learners interpret and respond to it. As Carless and Boud (2021) argue, feedback effectiveness ultimately depends on learners' sense-making processes rather than the feedback itself.

Importantly, these findings also have broader implications for the ongoing shift toward ESL-oriented instruction in Vietnam. Rather than suggesting that oral feedback in VSTEP preparation directly develops ESL-oriented communicative competence, the present findings indicate that balanced teacher oral feedback may support learners' perceived speaking development and confidence, both of which are valuable foundations for communication beyond examination settings. In this sense, effective feedback can be viewed as a pedagogical bridge between test preparation and broader communicative learning goals rather than as a direct cause of ESL-oriented competence. This includes promoting interaction, learner autonomy, and meaningful language use, as suggested by Ellis (2020) and Li (2010). Furthermore, the integration of AI-mediated feedback may complement teacher feedback by providing immediate and individualized input, although its effectiveness remains dependent on learners' engagement with such tools (Shi & Aryadoust, 2024; Godwin-Jones, 2022).

The findings suggest that effective oral feedback in VSTEP speaking classrooms requires a balanced combination of clarity, focus, appropriate timing, and emotional support. These dimensions were statistically associated with learners' perceptions of speaking development and confidence; however, the present study did not measure actual speaking performance or classroom behavior. Consequently, the pedagogical implications discussed here should be interpreted as being based on learners' reported perceptions rather than direct evidence of learning outcomes.

The findings suggest several practical implications for teachers working in VSTEP speaking preparation classes. First, oral feedback should be clear, concise, and directly linked to the VSTEP speaking assessment criteria so that learners understand exactly which aspects of their performance require improvement. Second, brief post-task feedback may be more effective than frequent interruptions during fluency-oriented speaking activities, thereby maintaining communicative flow while still providing opportunities for reflection. Third, teachers should

balance comments on learners' strengths with constructive suggestions for improvement in order to sustain motivation. Finally, repeated speaking errors should be addressed using supportive and encouraging language, as emotionally positive feedback may help maintain learners' confidence and willingness to communicate despite the pressures associated with high-stakes speaking assessments.

6. Conclusion

This study investigated learners' perceptions of teacher oral feedback in VSTEP speaking preparation classes at Lac Hong University, with particular attention to four key feedback dimensions: clarity, timing, focus, and affective tone. By examining how learners perceived these dimensions and their relationships with perceived speaking development and perceived confidence, the study sought to contribute to a better understanding of teacher oral feedback within a high-stakes speaking preparation context in Vietnam.

From an academic perspective, the study extends previous feedback research by focusing on learners' interpretations of teacher oral feedback rather than teachers' feedback practices alone. The findings suggest that the cognitive dimensions of feedback, particularly clarity and focus, were more strongly associated with learners' perceived speaking development, whereas affective tone was more closely associated with perceived confidence. These findings contribute to the growing body of research emphasizing the learner-centered nature of feedback and highlight the importance of considering both cognitive and affective dimensions when examining oral feedback in high-stakes speaking contexts.

From a pedagogical perspective, the findings suggest that teachers should provide oral feedback that is clear, focused, timely, and emotionally supportive. In VSTEP speaking preparation classes, brief post-task feedback, explicit links to VSTEP speaking assessment criteria, balanced comments on learners' strengths and weaknesses, delayed correction during fluency-focused activities, and supportive language when addressing repeated errors may help learners better understand speaking expectations while maintaining their confidence and willingness to communicate. Rather than suggesting that teacher oral feedback directly develops ESL-oriented communicative competence, the findings indicate that balanced oral feedback may support learners' perceived speaking development and confidence, both of which are valuable foundations for communication beyond examination settings.

Several limitations should be acknowledged. First, the study relied on a single self-report questionnaire administered to a relatively small sample of learners from one language center at one university, limiting the generalizability of the findings. Second, because the study collected only learners' self-reported perceptions, responses may have been influenced by self-report bias. Third, no qualitative data, classroom observations, teacher perspectives, or recordings of teacher oral feedback were collected to triangulate learners' responses. Finally, the study did not include objective measures of speaking performance; therefore, the reported relationships should be interpreted as associations between learners' perceptions rather than evidence of causal effects or actual improvements in speaking ability. Future studies could adopt mixed-methods designs combining questionnaires with classroom observations, semi-structured interviews, teacher feedback recordings, and analyses of learner uptake. Including objective measures such as pre- and post-speaking test scores would provide stronger evidence regarding the relationship between teacher oral feedback and speaking performance. Such approaches would also help explain why

learners perceive particular feedback dimensions as more useful or emotionally supportive within high-stakes speaking preparation contexts.

Overall, the present study contributes to the growing literature on teacher oral feedback by providing learner-centered evidence from a Vietnamese high-stakes VSTEP context, offering practical implications for improving speaking instruction while supporting broader communicative learning goals.

Acknowledgments

The author would like to thank Lac Hong University and all participating students for their valuable contributions to this research.

References

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Ellis, R. (2020). A short history of SLA: Where have we come from and where are we going? *Language Teaching, 54*(2), 190-205. https://doi.org/10.1017/S0261444820000038

Fulcher, G. (2024). *Practical language testing*. Routledge. https://doi.org/10.4324/9781003373629

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology, 26*(2), 5–24. https://doi.org/10.64152/10125/73474

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing, 30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60*(2), 309–365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*. https://doi.org/10.1111/j.1467-1770.1989.tb00423.x

Mäkipää, T. (2024). Upper secondary students' perceptions of feedback literacy in second language learning in Finland – A qualitative case study. *Teaching and Teacher Education, 143*, Article 104554. https://doi.org/10.1016/j.tate.2024.104554.

Doan, N., Pham, T., Pham, M., & Tran, K. (2018). English as an international language in Viet Nam: History and development. *Asian Englishes, 20*(2), 106–121. https://doi.org/10.1080/13488678.2018.1439324

Sato, M., & Loewen, S. (Eds.). (2019). *Evidence-based second language pedagogy: A collection of instructed second language acquisition studies*. Routledge. https://doi.org/10.4324/9781351190558

Stickler, U. (2022). *Technology and language teaching*. Cambridge University Press. https://doi.org/10.1017/9781108874403

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*. https://doi.org/10.1007/s10734-017-0220-3

Winstone, N., & Carless, D. (2019). *Designing effective feedback processes in higher education*. Routledge. https://doi.org/10.4324/9781351115940

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL, 36*(2), 187–209. https://doi.org/10.1017/S0958344023000265

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

#127 SHORT-FORM VIDEO PLATFORMS AS DIGITAL TOOLS FOR ESL VOCABULARY DEVELOPMENT: EVIDENCE FROM VIETNAMESE UNIVERSITY STUDENTS

Nguyen Thi Nuong[1], Phan Thi Thanh Tam[2], Nguyen Thanh Dung[3], Duong Thi Thu Uyen[4]

Abstract

This study investigates the role of short-form video platforms, particularly TikTok and Instagram Reels, in supporting English vocabulary learning among Vietnamese EFL university students. Based on Cognitive Theory of Multimedia Learning and micro-learning principles, the study examines the effects of engagement, repetition, multimedia elements, and distraction on vocabulary acquisition. A quantitative research design was employed, and data were collected from 220 university students through an online questionnaire. The data were analyzed using reliability analysis, Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), and Structural Equation Modeling (SEM). The findings indicate that repetition was the only statistically significant predictor of vocabulary acquisition ($\beta = .549$, $p = .012$), whereas engagement, multimedia elements, and distraction did not demonstrate statistically significant effects. The results suggest that repeated encounters with vocabulary items play a more important role in vocabulary learning than entertainment-oriented features or passive multimedia exposure. The study contributes to the growing literature on digital language learning and provides pedagogical implications for integrating short-form video platforms into technology-enhanced vocabulary instruction.

***Keywords*:** short-form video platforms, vocabulary acquisition, repetition, micro-learning, ESL, Vietnam, TikTok, Instagram Reels.

1. Introduction

English has become the dominant global lingua franca in education, business, science,

English has become the dominant global lingua franca in education, business, science, and international communication. In the Vietnamese context, English is generally regarded as a foreign language (EFL), although recent educational reforms have increasingly emphasized communicative competence and international integration. The growing importance of English proficiency has been institutionalized through national educational policies that promote English as a compulsory subject from primary school onward. Consequently, improving English language competence, particularly vocabulary knowledge, has become a major educational priority.

Vocabulary plays a fundamental role in second language acquisition because it forms the basis for listening, speaking, reading, and writing skills. Without sufficient vocabulary knowledge,

[1]*Nguyen Thi Nuong, Faculty of Foundation Studies and Foreign Languages, Academy of Policy and Development, Hanoi City, Viet Nam, Email: nuongnguyen@apd.edu.vn*
[2] *Faculty of Foreign Languages, University of Labour and Social Affairs, Hanoi city, Viet Nam.*
[3] *Department of General Education, National University of Arts Education*
[4] *Faculty of Foundation Studies and Foreign Languages, Academy of Policy and Development*

learners struggle to communicate effectively regardless of their grammatical competence. Despite its importance, vocabulary acquisition remains challenging for many English as a Foreign Language (EFL) learners due to traditional teaching approaches that rely heavily on memorization and decontextualized exercises. In recent years, digital technologies and social media platforms have transformed language learning environments. Among these technologies, short-form video platforms such as TikTok and Instagram Reels have become increasingly popular among university students. These platforms provide highly visual, interactive, and concise content that aligns with contemporary learners' media consumption habits. The growing popularity of short-form video content is particularly significant in Southeast Asia. According to Digital 2025 reports, Vietnamese users spend several hours each week consuming short-form videos on social media platforms. Beyond entertainment purposes, these platforms are increasingly used for informal learning activities, including English vocabulary development. The pedagogical potential of short-form videos can be explained through several theoretical perspectives. Mayer's Cognitive Theory of Multimedia Learning suggests that learners process information more effectively when verbal and visual information are presented simultaneously. Similarly, micro-learning theory emphasizes the effectiveness of short, focused learning units that reduce cognitive overload and increase learner engagement. Short-form videos integrate audio, subtitles, visual illustrations, and contextualized language use, thereby potentially supporting vocabulary acquisition in meaningful and memorable ways. Previous studies have shown that social media-based learning can enhance learner motivation, engagement, and vocabulary retention. However, existing research presents several limitations. First, many studies focus on a single platform rather than comparing multiple short-form video environments. Second, empirical research examining the direct relationship between specific video characteristics and vocabulary acquisition remains limited. Third, little research has investigated these relationships within the Vietnamese higher education context.

Although previous studies have highlighted the educational potential of social media platforms, most research has focused on learners' motivation, engagement, attitudes, or general platform use. Relatively little attention has been paid to identifying which specific characteristics of short-form video platforms are associated with vocabulary learning outcomes. In particular, the respective roles of engagement, repetition, multimedia elements, and distraction remain insufficiently understood. By examining these factors simultaneously within a single structural model, the present study contributes to a more nuanced understanding of how short-form video platforms may support English vocabulary learning among Vietnamese EFL university students.

To address these gaps, the present study investigates the impact of engagement, repetition, multimedia elements, and distraction on English vocabulary acquisition through TikTok and Instagram Reels among Vietnamese university students.

The study aims to answer the following research questions:

RQ1. *Which factors influence vocabulary acquisition through short-form video platforms?*

RQ2. *To what extent do these factors affect English vocabulary acquisition?*

RQ3. *Which factor exerts the strongest influence on vocabulary learning outcomes?*

The findings are expected to contribute both theoretically and practically to the field of digital language learning and technology-enhanced vocabulary instruction.

2. Literature Review

2.1. Social media and language learning

The rapid expansion of digital technologies has transformed how people communicate, interact, and access knowledge. Among these developments, social media platforms have become influential environments for information sharing, collaboration, and learning. Kaplan and Haenlein (2010) define social media as Internet-based applications built on Web 2.0 that enable the creation and exchange of user-generated content, while Boyd and Ellison (2007) describe social networking sites as web-based services that allow users to create profiles, establish connections, and interact within online communities. Their increasing accessibility has significantly influenced educational practices, particularly language learning.

Recent research has highlighted the pedagogical value of social media in second and foreign language education. These platforms promote communication, collaboration, learner autonomy, and authentic language exposure, supporting interactive and learner-centered learning (Manca & Ranieri, 2016). Kabilan et al. (2010) found that social networking environments enhance vocabulary development, writing fluency, and communicative competence through authentic interaction. From a sociocultural perspective, Vygotsky (1978) argues that social interaction is fundamental to cognitive development, and online communities provide opportunities for learners to negotiate meaning and co-construct knowledge (Reinhardt, 2019). These environments may also reduce language anxiety and encourage communication (Lee, 2014).

Social media has also been associated with greater learner motivation and engagement. Junco et al. (2011), for example, reported positive effects on student engagement and academic performance. Video-based content is particularly valuable because audiovisual input facilitates learning through multiple sensory channels (Mayer, 2009), improving listening comprehension, vocabulary acquisition, and pronunciation in authentic contexts (Vanderplank, 2016). Despite these benefits, social media may also create learning challenges. Excessive use can negatively affect academic performance and time management (Kirschner & Karpinski, 2010), while frequent digital multitasking may reduce sustained attention and cognitive performance (Rosen et al., 2013). These contrasting findings suggest that the relationship between social media use and language learning remains complex and context-dependent.

Consequently, greater attention should be given to examining how specific forms of social media content influence language learning. Among current digital trends, short-form video platforms such as TikTok and Instagram Reels have attracted increasing scholarly interest because of their rapid global growth and widespread popularity among younger learners.

2.2 Short-form videos and micro-learning

Short-form videos are concise audiovisual content, typically lasting from a few seconds to several minutes, and have gained widespread popularity through platforms such as TikTok, Instagram Reels, and YouTube Shorts. The educational potential of short-form videos can be attributed partly to their concise format and engaging audiovisual features. Research suggests that concise and visually engaging instructional videos can enhance learner attention and engagement. Guo et al. (2014) found that shorter videos achieved higher completion rates and sustained learner attention, whereas Bétrancourt & Benetos (2018) emphasized that the effectiveness of instructional videos depends on how video content and design interact with learners' cognitive and perceptual

processes. The effectiveness of short-form videos is also supported by Mayer's (2009) Cognitive Theory of Multimedia Learning, which suggests that learners process verbal and visual information through dual channels. Accordingly, multimedia elements such as narration, subtitles, visuals, and contextual demonstrations can enhance comprehension and memory retention. Likewise, Cognitive Load Theory (Sweller, 1988) proposes that presenting information in manageable segments reduces cognitive overload, although excessive pacing or visual complexity may hinder learning (Paas & Sweller, 2014).

The growing popularity of TikTok and Instagram Reels has encouraged research on their educational applications. Escamilla-Fajardo et al. (2021) suggest that TikTok promotes active and collaborative learning through its interactive features. Similarly, Dizon (2021) found that learners held favorable perceptions of video streaming for informal foreign language learning, highlighting its potential to support out-of-class language learning. In addition, captioned audiovisual input can support learners' comprehension and language development (Vanderplank, 2016).

2.3. Vocabulary acquisition in digital environments

Vocabulary knowledge is a fundamental component of language proficiency because it underpins learners' ability to comprehend and produce language across listening, speaking, reading, and writing (Nation, 2001; Schmitt, 2010). Contemporary research suggests that vocabulary learning is a gradual process influenced by repeated exposure, meaningful input, cognitive engagement, and opportunities for retrieval and use (Webb & Nation, 2017). Based on previous research, the present study examines four factors influencing vocabulary learning through short-form video platforms: repetition, multimedia elements, engagement, and distraction.

First, repetition is widely recognized as a key mechanism for vocabulary acquisition. Nation (2001) argues that repeated encounters strengthen form meaning connections, while Webb (2007) suggests that multiple exposures improve learners' ability to recognize, recall, and use vocabulary. Similar findings have been reported by Pellicer-Sánchez & Schmitt (2010), Waring and Takaki (2003), and Horst et al. (1998). However, most previous studies were conducted in reading-based or classroom settings, leaving the effectiveness of algorithm-driven repeated exposure on short-form video platforms insufficiently explored.

Second, multimedia elements play an important role in digital vocabulary learning. According to the Cognitive Theory of Multimedia Learning (Mayer, 2009) and Dual Coding Theory (Paivio, 1991), combining verbal and visual information enhances learning. In language education, subtitles, images, narration, and contextual demonstrations support vocabulary recognition and comprehension (Vanderplank, 2016). Previous studies have reported positive effects of audiovisual input and captioned videos on incidental vocabulary learning (Montero Perez et al., 2014; Peters & Webb, 2018). Nevertheless, it remains unclear whether multimedia directly improves vocabulary acquisition or mainly increases learners' attention and interest.

Third, engagement reflects learners' behavioral, emotional, and cognitive involvement in learning (Fredricks et al., 2004). Greater engagement is generally associated with higher attention, persistence, and learning effectiveness (Skinner & Pitzer, 2012). However, prolonged engagement does not necessarily result in meaningful vocabulary processing, making its contribution to vocabulary learning uncertain.

Finally, despite their educational benefits, short-form video platforms may also create distractions that interfere with learning. Cognitive Load Theory suggests that excessive stimulation can overload working memory and reduce learning efficiency (Sweller et al., 2011). Rapid transitions, background music, algorithmic recommendations, and entertainment-oriented content may divert learners' attention, while digital multitasking has been shown to impair concentration and academic performance (Ophir et al., 2009; Rosen et al., 2013). At the same time, some evidence suggests that learners can adapt to highly stimulating digital environments, leaving the relationship between distraction and vocabulary learning inconclusive.

Overall, the literature indicates that repetition, multimedia elements, engagement, and distraction influence vocabulary learning through different cognitive and behavioral mechanisms. However, evidence regarding their relative contributions in short-form video environments remains limited and inconsistent, highlighting the need for further investigation into how these factors shape vocabulary learning on platforms such as TikTok and Instagram Reels.

2.4. Research model and hypotheses

Drawing upon the theoretical foundations discussed in the preceding literature review, this study proposes a conceptual research model to examine the factors influencing English vocabulary acquisition through short-form video platforms, particularly TikTok and Instagram Reels. The model is grounded primarily in Mayer's (2009) Cognitive Theory of Multimedia Learning, Sweller's (1988) Cognitive Load Theory, and vocabulary acquisition research emphasizing repetition, multimodal input, and learner engagement (Nation, 2001; Webb & Nation, 2017).

The proposed framework conceptualizes vocabulary acquisition as the dependent construct influenced by four independent variables: engagement (ENG), repetition (REP), multimedia elements (MUL), and distraction (DIS). These constructs were selected because they have been consistently identified in previous research on digital learning environments, multimedia-assisted language learning, and social media-based learning. Although engagement, repetition, and multimedia elements may co-occur in authentic short-form video environments, they represent conceptually distinct dimensions of the learning experience. For example, videos containing rich multimedia features may attract greater learner engagement and encourage repeated viewing. Nevertheless, engagement refers to learners' behavioral, emotional, and cognitive involvement while interacting with content (Fredricks et al., 2004), repetition concerns the frequency with which learners encounter and revisit vocabulary items, and multimedia elements refer to the visual, auditory, and textual features that support meaning construction. Although these constructs may be related in practice, they were operationalized as theoretically distinct dimensions and measured using separate questionnaire items derived from their respective conceptual definitions.

Engagement refers to the degree of learners' behavioral, emotional, and cognitive involvement while interacting with short-form video content (Fredricks et al., 2004). Previous studies suggest that engaged learners are more likely to sustain attention, invest effort, and participate actively in learning activities, which may contribute to more effective vocabulary acquisition (Junco et al., 2011; Dörnyei & Ushioda, 2011). In the context of TikTok and Instagram Reels, highly interactive and visually appealing videos may encourage learners to spend more time engaging with English vocabulary content. Therefore, engagement is expected to positively influence vocabulary acquisition.

Repetition represents the frequency with which learners encounter and review lexical items through repeated exposure to short-form videos. Vocabulary acquisition research consistently identifies repetition as a critical factor supporting long-term memory retention and lexical consolidation (Nation, 2001; Webb, 2007). Repeated encounters with vocabulary items strengthen form meaning connections and facilitate retrieval from long-term memory (Pellicer-Sánchez & Schmitt, 2010). Given that algorithm-driven social media platforms frequently recommend similar content, learners may repeatedly encounter target vocabulary across different contexts, thereby reinforcing vocabulary learning. Consequently, repetition is hypothesized to positively influence vocabulary acquisition.

Multimedia elements refer to the integration of visual illustrations, audio narration, subtitles, animations, and contextual demonstrations within short-form videos. According to Cognitive Theory of Multimedia Learning (Mayer, 2009) and Dual Coding Theory (Paivio, 1991), combining verbal and visual input enhances cognitive processing and information retention. Previous studies have shown that subtitles and audiovisual input can facilitate incidental vocabulary acquisition and listening comprehension (Montero Perez et al., 2014; Peters & Webb, 2018). Therefore, multimedia elements are expected to facilitate vocabulary acquisition by supporting comprehension and retention through multiple channels of information processing.

In the present study, distraction refers to learners' perceived experience of cognitive interference while using short-form video platforms. It does not measure actual attentional loss or objectively observed multitasking behavior. Cognitive Load Theory suggests that excessive information-processing demands may overload working memory and negatively affect learning effectiveness (Sweller, 1988). Similarly, studies on digital multitasking indicate that frequent distractions may reduce concentration and impair academic performance (Ophir et al., 2009; Rosen et al., 2013). Accordingly, perceived distraction is hypothesized to negatively influence vocabulary acquisition. Based on the theoretical and empirical foundations discussed above, the study proposes the following hypotheses:

H1: Engagement (ENG) positively affects vocabulary acquisition (VOC).

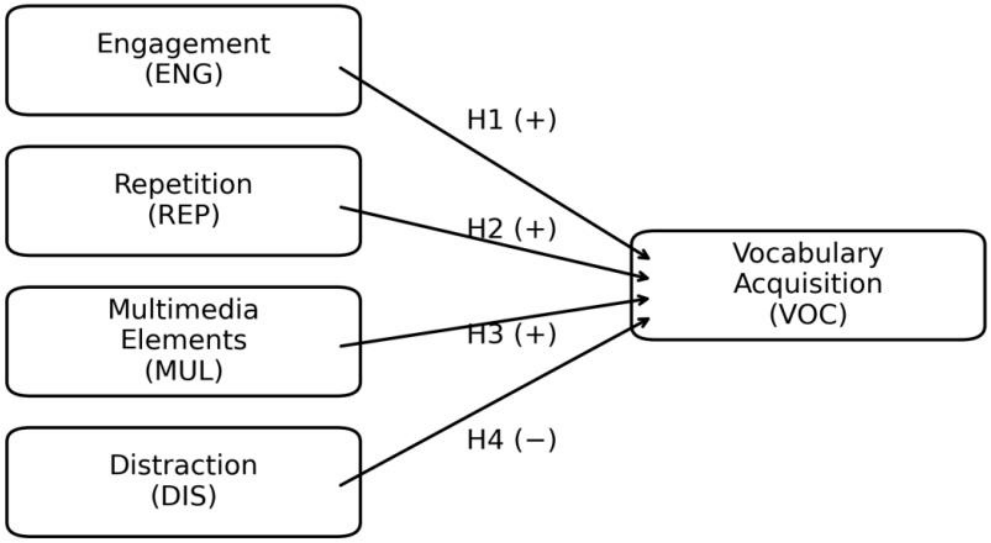
H2: Repetition (REP) positively affects vocabulary acquisition (VOC).

H3: Multimedia elements (MUL) positively affect vocabulary acquisition (VOC).

H4: Distraction (DIS) negatively affects vocabulary acquisition (VOC).

Figure 1

Research model and hypotheses



3. Methodology

3.1. Research design

This study employed a quantitative research design to investigate the relationships between short-form video characteristics and English vocabulary acquisition among Vietnamese university students. A quantitative approach was considered appropriate because the study aimed to examine the magnitude and direction of relationships among multiple variables and to statistically test the proposed research hypotheses. The research framework was developed based on theories of multimedia learning, cognitive load, and vocabulary acquisition in digital learning environments. Specifically, the study examined the effects of four independent variables engagement (ENG), repetition (REP), multimedia elements (MUL), and distraction (DIS) on the dependent variable, vocabulary acquisition (VOC). A cross-sectional survey design was adopted to collect empirical data from participants at a single point in time. This approach enabled the researchers to efficiently gather data regarding learners' experiences and perceptions of using TikTok and Instagram Reels for English vocabulary learning. The study further used Structural Equation Modeling (SEM) as the primary analytical technique to evaluate the relationships among the constructs and validate the proposed conceptual framework.

3.2. Participants

The participants in this study consisted of 220 Vietnamese university students who regularly used short-form video platforms, particularly TikTok and Instagram Reels, for English learning purposes. The sample size of 220 participants was considered adequate for Structural Equation Modeling (SEM). According to Hair et al. (2010), a sample size exceeding 200 observations is generally acceptable for covariance-based SEM models with moderate complexity. In addition, the sample satisfied the commonly recommended ratio of at least 10 respondents per estimated parameter, thereby providing sufficient statistical power for the analyses conducted in this study. The participants were recruited using convenience sampling through online learning communities and university-related social media groups.

The sample primarily represented university-aged learners who actively engaged in self-directed English learning through digital platforms. Most participants reported intermediate levels of English proficiency, indicating that they possessed sufficient linguistic competence to evaluate the effectiveness of short-form video content for vocabulary acquisition. In terms of gender distribution, female students constituted the majority of the sample. The selection of university students as the target population was considered appropriate because young adult learners are among the most active users of social media platforms and frequently engage with short-form video content in both educational and entertainment contexts. Furthermore, university students are generally familiar with mobile-assisted language learning practices, making them suitable participants for investigating digital vocabulary learning behaviors.

Prior to data collection, participants were informed about the purpose of the study and assured that their responses would remain anonymous and confidential. Participation in the survey was entirely voluntary.

3.3. Research instrument

Data were collected through an online questionnaire administered via Google Forms. The questionnaire was designed based on established measurement scales from previous studies on digital learning, multimedia learning, learner engagement, and vocabulary acquisition. To ensure contextual appropriateness, several items were adapted and modified to align with the Vietnamese higher education context and the specific characteristics of short-form video platforms.

The questionnaire consisted of two main sections. The first section collected demographic information, including gender, age, English proficiency level, learning experience, and the frequency of using TikTok and Instagram Reels for English learning purposes. The second section measured the research constructs, namely engagement (ENG), repetition (REP), multimedia elements (MUL), distraction (DIS), and vocabulary acquisition (VOC). Each construct was operationalized through multiple observed variables adapted from prior literature. All questionnaire items were measured using a five-point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). The Likert-scale format was selected because it allows respondents to express varying degrees of agreement and is widely used in educational and behavioral research. The engagement construct assessed learners’ levels of attention, interest, and interaction while using short-form videos. The repetition construct measured the extent to which learners repeatedly encountered vocabulary items across videos. Multimedia elements examined learners’ perceptions of subtitles, visuals, audio, and contextual illustrations in facilitating vocabulary learning. Distraction assessed the extent to which rapid pacing, entertainment-oriented content, and excessive stimulation interfered with learning processes. Finally, vocabulary acquisition evaluated learners’ perceived improvement in vocabulary retention, understanding, and usage.

To ensure content validity, the questionnaire items were reviewed and refined based on relevant theoretical frameworks and empirical studies in second language acquisition and digital learning research.

3.4. Data analysis

The collected data were analyzed using IBM SPSS Statistics 26 and IBM SPSS AMOS 26. Prior to analysis, the dataset was screened for completeness and consistency. Descriptive statistics were used to summarize participants' demographic characteristics. Cronbach's Alpha was employed to assess the internal consistency of the measurement scales, while Exploratory Factor Analysis (EFA) was conducted to examine the underlying factor structure and construct validity. The suitability of the data for factor analysis was evaluated using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity.

Multiple regression analysis was initially performed to examine the predictive relationships between the independent variables and vocabulary acquisition. Subsequently, Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were conducted using AMOS 26. CFA was used to assess the reliability and validity of the measurement model, whereas SEM was employed to test the hypothesized relationships among the latent constructs. SEM was selected because the study involved latent variables measured by multiple observed indicators and allowed the simultaneous assessment of both measurement and structural models while accounting for measurement error. Model fit was evaluated using χ^2/df, CFI, TLI, and RMSEA.

4. Results and discussion

4.1. Descriptive statistics

Table 1

Frequency distribution of the research sample (N=220)

Variable	Category	Frequency (N)	%	Valid %	Cumulative %
Gender	Male	49	22.3	22.3	22.3
	Female	171	77.7	77.7	100.0
Age	Group 1	6	2.7	2.7	2.7
	Group 2	206	93.6	93.6	96.4
	Group 3	8	3.6	3.6	100.0
English proficiency	Level 1	55	25.0	25.0	25.0
	Level 2	150	68.2	68.2	93.2
	Level 3	15	6.8	6.8	100.0
English learning experience	Level 1	17	7.7	7.7	7.7
	Level 2	28	12.7	12.7	20.5
	Level 3	33	15.0	15.0	35.5
	Level 4	142	64.5	64.5	100.0
Self - learning	Level 1	13	5.9	5.9	5.9
	Level 2	26	11.8	11.8	17.7

Level 3	136	61.8	61.8	79.5
Level 4	38	17.3	17.3	96.8
Level 5	7	3.2	3.2	100.0

Source: Analysis results obtained from SPSS 26

After collecting data through the online survey, the research team obtained a total of 220 valid responses, which were included in the analysis. The research sample is described based on several fundamental demographic characteristics, including gender, age, English proficiency, English learning experience, and the participants' level of self-study.

4.2. Reliability analysis

Prior to hypothesis testing, the internal consistency and reliability of the measurement scales were evaluated using Cronbach's Alpha coefficients. According to Nunnally and Bernstein (1994), a Cronbach's Alpha value above .70 indicates acceptable reliability for social science research. The results demonstrated that all constructs achieved satisfactory reliability levels, confirming the consistency of the measurement items used in this study.

Specifically, the engagement construct (ENG) obtained a Cronbach's Alpha coefficient of .819, indicating good internal consistency among the observed variables measuring learners' interaction and motivation when using short-form videos. The repetition construct (REP) achieved an Alpha coefficient of .815, suggesting that the items consistently reflected repeated lexical exposure within short-form video environments. The multimedia elements construct (MUL) showed a relatively high reliability level with a Cronbach's Alpha value of .865, indicating strong consistency among items related to subtitles, visuals, audio, and contextual illustrations. Similarly, distraction (DIS) demonstrated the highest reliability among the independent variables with an Alpha coefficient of .878, reflecting strong coherence among the items measuring cognitive interference and overstimulation. The dependent variable, vocabulary acquisition (VOC), achieved a Cronbach's Alpha value of .830, indicating satisfactory reliability in measuring learners' perceived vocabulary development.

Overall, all scales exceeded the recommended reliability threshold, and no observed variables were removed from the analysis. These findings confirm that the measurement model possesses adequate internal consistency and is suitable for subsequent factor analysis and hypothesis testing.

Table 2.

Reliability analysis of measurement scales

Construct	**Number of items**	**Cronbach's Alpha**	**Reliability level**
Engagement (ENG)	2	.819	Good
Repetition (REP)	2	.815	Good
Multimedia Elements (MUL)	3	.865	Very Good
Distraction (DIS)	2	.878	Excellent
Vocabulary Acquisition (VOC)	3	.830	Good

Source: Authors' calculation based on SPSS output.

4.3. Exploratory factor analysis

To examine the construct validity and underlying factor structure of the measurement model, Exploratory Factor Analysis (EFA) was conducted using Principal Component Analysis with Varimax rotation. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy yielded a value of .928, substantially exceeding the recommended threshold of .50 (Hair et al., 2010). This result indicates that the dataset was highly appropriate for factor analysis. Furthermore, Bartlett's Test of Sphericity was statistically significant ($p < .001$), confirming that the observed variables were sufficiently correlated to justify factor extraction.

The EFA results indicated that the extracted factors explained 66.451% of the total variance, demonstrating acceptable construct validity for the proposed measurement model. In social science research, a total variance explained value above 50% is generally considered acceptable, indicating that the factor structure adequately represents the observed data. The extracted factor structure generally aligned with the proposed theoretical framework. The independent variables including engagement, repetition, multimedia elements, and distraction loaded strongly onto the primary latent dimension, while vocabulary acquisition emerged as a separate dependent construct. The factor loadings exceeded the acceptable threshold, supporting the convergent validity of the measurement scales.

Overall, the EFA results confirm that the observed variables appropriately represent the underlying theoretical constructs and provide empirical support for the proposed research model.

Table 3

KMO and Bartlett's Test

Measure	Value
Kaiser–Meyer–Olkin (KMO)	.928
Bartlett's Test of Sphericity	Significant
Sig.	< .001

Source: Authors' calculation based on SPSS output.

Table 4.

Total variance explained

Factor	Eigenvalue	% of Variance	Cumulative %
Factor 1	7.214	48.093	48.093
Factor 2	2.754	18.358	66.451

Source: Authors' calculation based on SPSS output.

4.4. Confirmatory factor analysis

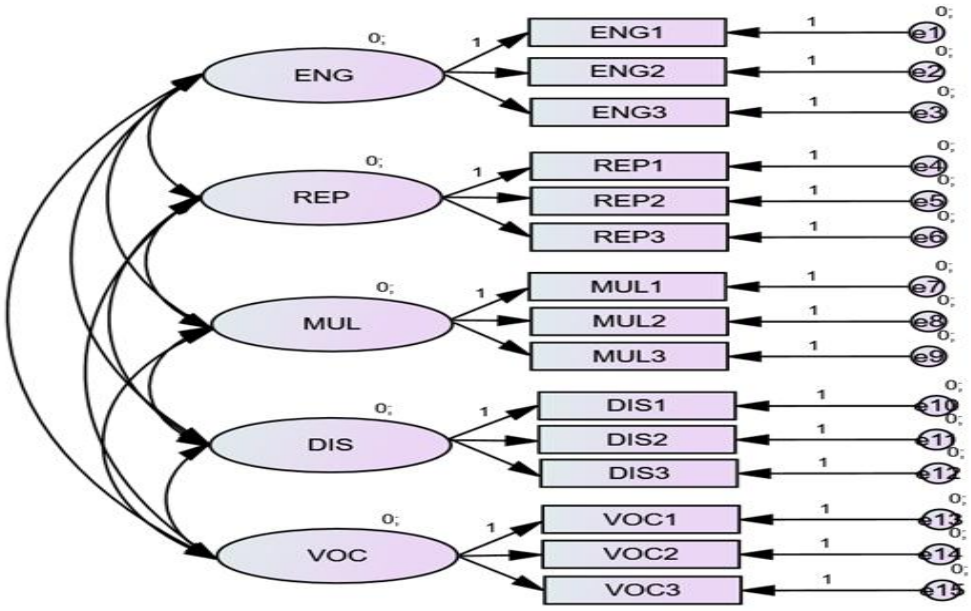
4.4.1. Objective

Confirmatory factor analysis (CFA) was conducted to validate the measurement model, examine the goodness-of-fit between the proposed model and the empirical data, and assess the convergent and discriminant validity of the constructs. Following the reliability analysis and Exploratory Factor Analysis (EFA), the CFA included five latent constructs: Engagement (ENG), Repetition (REP), Multimedia Elements (MUL), Distraction (DIS), and Vocabulary Acquisition (VOC), represented by 15 observed variables.

4.4.2. Measurement model

Figure 2

Confirmatory factor analysis (CFA) measurement model



The measurement model demonstrated acceptable fit indices ($\chi^2/df = 2.11$, CFI = .931, TLI = .918, RMSEA = .061), indicating a satisfactory fit between the proposed model and the observed data (Hair et al., 2010). The CFA results showed that all standardized factor loadings were statistically significant ($p < .001$) and exceeded the recommended threshold of 0.50, ranging from 0.718 to 0.884 (Hair et al., 2010). These findings indicate that all observed variables adequately represented their corresponding latent constructs and provide evidence of satisfactory convergent validity. Furthermore, the reliability analysis reported in the previous section showed that all constructs achieved Cronbach's Alpha values above the recommended threshold of .70, indicating acceptable internal consistency (Nunnally & Bernstein, 1994). Taken together, the CFA results, reliability analysis, and acceptable model fit indices support the adequacy of the proposed measurement model for subsequent structural analysis.

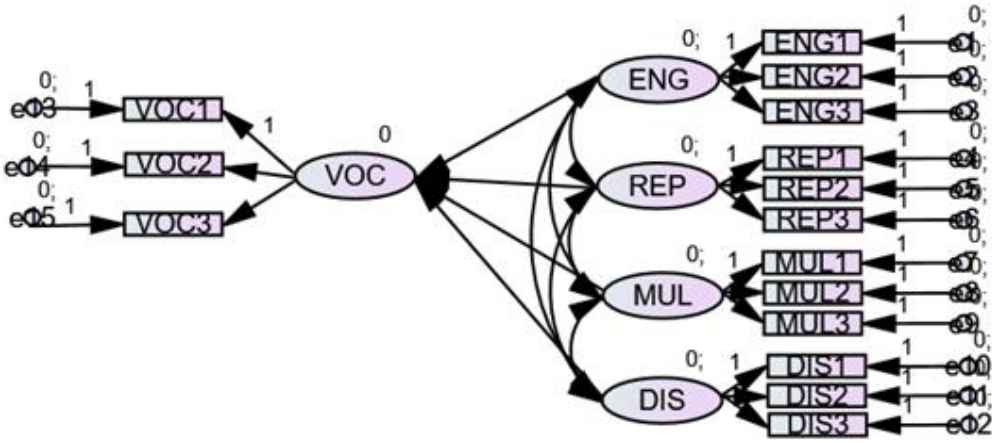
The mean values of the observed variables ranged from 3.436 to 3.986, reflecting a moderately high level of agreement among respondents. The Multimedia Elements (MUL) construct obtained the highest mean scores, suggesting that learners highly appreciated the role of

multimedia features in supporting vocabulary learning. By contrast, the Distraction (DIS) construct recorded comparatively lower mean scores, indicating that distraction was generally perceived at a moderate level.

4.5. Structural equation modeling (SEM) results

Figure 3

Structural equation model with standardized estimates



The Structural Equation Modeling (SEM) analysis was conducted to examine the effects of Engagement (ENG), Repetition (REP), Multimedia Elements (MUL), and Distraction (DIS) on Vocabulary Acquisition (VOC). The results indicate that engagement (ENG) did not significantly affect vocabulary acquisition (β = .118, p = .231). This finding suggests that learners' interaction and engagement with short-form video content alone may not necessarily lead to improved vocabulary acquisition. In contrast, repetition (REP) demonstrated a positive and statistically significant effect on vocabulary acquisition (β = .549, p = .012). This result highlights the critical role of repeated lexical exposure in reinforcing vocabulary retention and enhancing learning effectiveness. Similarly, multimedia elements (MUL) did not significantly influence vocabulary acquisition (β = .093, p = .187). Although multimedia features may increase learner attention and engagement, they may not independently produce substantial vocabulary gains without repeated exposure and meaningful processing. Finally, distraction (DIS) did not demonstrate a statistically

significant effect on vocabulary acquisition (β = –.074, p = .294). This finding may indicate that learners are capable of adapting to high-stimulation digital environments and selectively focusing on relevant linguistic information during learning activities.

Overall, among the four proposed predictors, repetition emerged as the only statistically significant factor influencing vocabulary acquisition in short-form video learning environments.

4.6. Regression and structural equation modeling (SEM) results

To examine the relationships between repetition, multimedia elements, engagement, distraction, and learners' perceived vocabulary acquisition, multiple regression analysis was initially conducted, followed by Structural Equation Modeling (SEM) to validate the findings while accounting for measurement error and simultaneously evaluating both the measurement and structural models. The regression results indicated that the proposed model demonstrated good explanatory power, accounting for 64.9% of the variance in learners' perceived vocabulary acquisition (Adjusted R^2 = .649). This finding suggests that the four predictor variables collectively provide a substantial explanation of vocabulary learning within short-form video environments.

Among the independent variables, repetition (REP) emerged as the strongest and only statistically significant predictor of perceived vocabulary acquisition. The regression coefficient was positive and significant (β = .372, p < .001), indicating that repeated exposure to vocabulary through short-form videos substantially enhanced learners' vocabulary acquisition. This result was further supported by the SEM analysis, in which the path from repetition to vocabulary acquisition remained statistically significant (β = .549, p = .012), confirming repetition as the most influential factor in the proposed structural model. By contrast, engagement (ENG), multimedia elements (MUL), and distraction (DIS) did not demonstrate statistically significant effects on perceived vocabulary acquisition. Although these constructs showed directional relationships with the dependent variable, their path coefficients did not reach statistical significance in the final structural model.

These findings suggest that repeated encounters with vocabulary items play a more decisive role in vocabulary learning than entertainment-oriented features, visual attractiveness, or passive interaction with short-form video content.

Table 5

Multiple regression results

Predictor	Standardized Beta (β)	t-value	Sig.
Engagement (ENG)	.104	1.421	.157
Repetition (REP)	.372	5.884	< .001
Multimedia Elements (MUL)	.087	1.238	.217
Distraction (DIS)	-.061	-0.942	.347

Table 6

Structural Equation Modeling (SEM) Results

Structural path	Estimate	p-value	Result
ENG → VOC	.118	.231	Not Supported
REP → VOC	.549	.012	Supported
MUL → VOC	.093	.187	Not Supported
DIS → VOC	-.074	.294	Not Supported

Source: Authors' calculation based on SEM analysis.

5. Discussion

5.1 Interpretation of findings

The findings of this study provide important insights into learners' perceived vocabulary acquisition through short-form video platforms such as TikTok and Instagram Reels. Among the four examined predictors, repetition emerged as the only statistically significant factor influencing vocabulary acquisition, highlighting the central role of repeated lexical exposure in digital language learning environments.

The significant influence of repetition is consistent with previous research in second language acquisition emphasizing the importance of repeated encounters with vocabulary items for long-term retention and lexical consolidation (Nation, 2001; Webb, 2007; Webb & Nation, 2017). From a cognitive perspective, repeated exposure strengthens form-meaning connections and facilitates the transfer of lexical information from working memory into long-term memory storage. Within short-form video environments, algorithm-driven recommendation systems may unintentionally support this process by repeatedly exposing learners to similar lexical items, expressions, and thematic content across multiple videos. The findings further suggest that repetition may play a more decisive role in vocabulary learning than entertainment-oriented or visually attractive characteristics of short-form videos. Although social media platforms are commonly associated with highly stimulating multimedia features, the present study demonstrates that repeated lexical exposure, rather than content richness alone, functions as the primary mechanism supporting vocabulary acquisition. This result reinforces the argument that frequency of exposure remains a fundamental determinant of vocabulary development regardless of technological learning environments. Contrary to the proposed hypotheses, engagement and multimedia elements did not significantly predict vocabulary acquisition. One possible explanation is that visually attractive content and interactive features may increase learners' attention and enjoyment without necessarily promoting deeper cognitive processing or durable vocabulary retention. Learners may passively consume entertaining content without systematically focusing on lexical learning. Consequently, while multimedia features may facilitate short-term engagement, they may not independently produce measurable vocabulary gains unless accompanied by repeated and meaningful exposure. Similarly, the non-significant effect of engagement indicates that increased interaction or enjoyment does not automatically translate into improved vocabulary acquisition. Although previous studies have linked learner engagement to stronger academic performance and motivation (Fredricks et al., 2004; Junco et al., 2011), the present findings suggest that engagement alone may be insufficient to generate substantial vocabulary learning outcomes in informal digital learning contexts. Interestingly, distraction did

not exert a statistically significant negative effect on vocabulary acquisition. This finding may indicate that contemporary learners have become increasingly accustomed to high-stimulation digital environments and may possess the ability to selectively process relevant linguistic information despite rapid pacing, background music, and entertainment-oriented content. In other words, learners may develop adaptive cognitive strategies that enable them to extract useful vocabulary input even within highly dynamic social media environments. Nevertheless, the absence of statistical significance should not be interpreted as evidence that distraction has no effect on learning. Rather, it suggests that within the present sample and research context, distraction did not substantially hinder vocabulary acquisition outcomes.

Overall, the findings indicate that effective vocabulary learning through short-form video platforms depends less on entertainment value or multimedia complexity and more on structured repetition and repeated lexical exposure. The study therefore highlights the pedagogical importance of repetition as a cognitive reinforcement mechanism in digital vocabulary learning environments and suggests that short-form videos can support vocabulary acquisition when designed to facilitate repeated encounters with target lexical items.

5.2. Theoretical and practical contributions

This study makes several important contributions to the existing literature on digital language learning, vocabulary acquisition, and social media-assisted education. First, the study extends current research on short-form video platforms by examining the pedagogical mechanisms underlying vocabulary acquisition within algorithm-driven social media environments such as TikTok and Instagram Reels. While previous studies have frequently emphasized learner motivation, engagement, and entertainment value in social media learning, the present study demonstrates that repeated lexical exposure functions as the primary mechanism supporting vocabulary development in short-form video contexts. In doing so, the study shifts scholarly attention from surface-level engagement toward deeper cognitive processes associated with vocabulary retention and lexical consolidation.

Second, the study contributes to vocabulary acquisition theory by reaffirming the central role of repetition in informal digital learning environments. Although multimedia learning research often highlights the benefits of audiovisual input and interactive media, the findings suggest that multimedia richness alone may not be sufficient to produce substantial vocabulary gains without continuous lexical recycling and repeated encounters with target vocabulary items. The study therefore provides a more nuanced understanding of how learners process language input in highly stimulating digital environments and highlights repetition as a fundamental pedagogical principle that remains effective even within contemporary social media ecosystems.

Third, the study contributes contextual evidence from the Vietnamese higher education context, which remains relatively underrepresented in current research on short-form video-assisted language learning. By focusing specifically on Vietnamese university students, the study broadens the geographical and educational scope of research on social media-based vocabulary

acquisition and provides empirical insights into how learners in emerging digital learning environments interact with short-form educational content.

From a practical perspective, the findings offer meaningful implications for educators, instructional designers, and digital content creators. The study suggests that vocabulary-oriented short-form videos should prioritize repeated exposure to target lexical items through subtitles, thematic recycling, contextual repetition, and spaced reinforcement strategies rather than relying primarily on entertainment-oriented features or excessive multimedia stimulation. Furthermore, teachers may integrate short-form video platforms into supplementary vocabulary instruction to extend learners' exposure to authentic English input beyond formal classroom settings and promote continuous vocabulary reinforcement through mobile-assisted learning practices.

6. Conclusion

This study investigated the effectiveness of short-form video platforms, particularly TikTok and Instagram Reels, in supporting English vocabulary acquisition among Vietnamese university students. Drawing upon the Cognitive Theory of Multimedia Learning, Cognitive Load Theory, and micro-learning principles, the study examined the effects of engagement, repetition, multimedia elements, and distraction on vocabulary acquisition through a quantitative research design employing Structural Equation Modeling (SEM). The findings revealed that repetition was the only statistically significant predictor of vocabulary acquisition, whereas engagement, multimedia elements, and distraction did not demonstrate significant effects.

The findings suggest that effective vocabulary learning in short-form video environments depends less on entertainment-oriented interaction or multimedia complexity and more on repeated lexical exposure and continuous vocabulary reinforcement. The study therefore highlights the pedagogical importance of repetition as a key mechanism supporting vocabulary retention in contemporary digital learning environments. From a practical perspective, educators and digital content creators should prioritize subtitles, contextual recycling, thematic consistency, and repeated encounters with target vocabulary items when designing vocabulary-oriented short-form video content. Despite its contributions, the study has several limitations. First, the research relied on self-reported questionnaire data collected from a relatively limited sample of Vietnamese university students, which may restrict the generalizability of the findings to other educational and cultural contexts. Second, the cross-sectional research design did not permit the examination of long-term vocabulary retention or causal relationships over time. Third, the study focused primarily on learners' perceptions rather than objective measures of vocabulary performance. In addition, the use of convenience sampling through online learning communities may have introduced self-selection bias, as students with greater interest in English learning or short-form video platforms were more likely to participate. The sample was also predominantly female, which may limit the representativeness of the findings across the broader university student population. Furthermore, several constructs, including engagement, repetition, and distraction, were measured using only two observed indicators. Although the reliability coefficients were acceptable, these measures may not fully capture the complexity of the underlying constructs. Finally, engagement, repetition, and multimedia elements may partially overlap in authentic short-form video learning environments, despite being operationalized as

theoretically distinct constructs in this study. Future studies may therefore employ longitudinal or experimental approaches, incorporate vocabulary testing instruments, and compare multiple social media platforms across diverse learner populations.

Overall, the study demonstrates that short-form video platforms possess considerable potential for supporting vocabulary acquisition when they facilitate repeated and meaningful exposure to target lexical items. As social media increasingly becomes integrated into language learning practices, understanding the cognitive and pedagogical mechanisms underlying digital vocabulary learning will remain an important area for future research.

References

Bétrancourt, M., & Benetos, K. (2018). Why and when does instructional video facilitate learning? A commentary to the special issue "Developments and trends in learning with instructional video". *Computers in Human Behavior, 89*, 471-475. https://doi.org/10.1016/j.chb.2018.08.035

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Dizon, G. (2021). Subscription video streaming for informal foreign language learning: Japanese EFL students' practices and perceptions. *TESOL Journal, 12*(2), 1–20.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Routledge.

Escamilla-Fajardo, P., Alguacil, M., & López-Carril, S. (2021). Incorporating TikTok in higher education: Pedagogical perspectives from a corporal expression sport sciences course. *Journal of Hospitality, Leisure, Sport & Tourism Education, 28*, 100302. https://doi.org/10.1016/j.jhlste.2021.100302

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of evidence. *Review of Educational Research, 74*(1), 59-109. https://doi.org/10.3102/00346543074001059

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning at Scale Conference* (pp. 41-50). ACM. https://doi.org/10.1145/2556325.2566239

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Pearson.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207-223.

Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. *Journal of Computer Assisted Learning, 27*(2), 119-132. https://doi.org/10.1111/j.1365-2729.2010.00387.x

Kabilan, M. K., Ahmad, N., & Abidin, M. J. Z. (2010). Facebook: An online environment for learning of English in institutions of higher education? *Internet and Higher Education, 13*(4), 179-187. https://doi.org/10.1016/j.iheduc.2010.07.003

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Kirschner, P. A., & Karpinski, A. C. (2010). Facebook and academic performance. *Computers in Human Behavior, 26*(6), 1237-1245. https://doi.org/10.1016/j.chb.2010.03.024

Lee, L. (2014). Digital news stories: Building language learners' content knowledge and speaking skills. *Foreign Language Annals, 47*(2), 338-356. https://doi.org/10.1111/flan.12084

Manca, S., & Ranieri, M. (2016). Facebook and the others: Potentials and obstacles of social media for teaching in higher education. *Computers & Education, 95*, 216-230. https://doi.org/10.1016/j.compedu.2016.01.012

Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology, 18*(1), 118-141.

Nation, I. S. P. (2001). Learning vocabulary in another language. Cambridge University Press.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences, 106*(37), 15583–15587. https://doi.org/10.1073/pnas.0903620106

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 27-42). Cambridge University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287. https://doi.org/10.1037/h0084295

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel. *Reading in a Foreign Language, 22*(1), 31-55.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition, 40*(3), 551-577. https://doi.org/10.1017/S0272263117000407

Reinhardt, J. (2019). Social media in second and foreign language teaching and learning: Blogs, wikis, and social networking. *Language Teaching, 52*(1), 1-39. https://doi.org/10.1017/S0261444818000356

Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. *Computers in Human Behavior, 29*(3), 948-958. https://doi.org/10.1016/j.chb.2012.12.001

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement. In S. Christenson et al. (Eds.), *Handbook of research on student engagement* (pp. 21-44). Springer.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching*. Palgrave Macmillan.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*(2), 130-163.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46-65. https://doi.org/10.1093/applin/aml048

Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.

Bionote

Nguyen Thi Nuong holds a PhD in Linguistics and is currently a lecturer at the Academy of Policy and Development, Viet Nam. She has over 18 years of experience in teaching, academic management, and professional development training for primary school teachers. Her teaching areas include General English, ESP, Business English, translation, communicative skills, linguistics, and intercultural communication. Her researches focus on applied linguistics, technology-enhanced language learning, digital transformation in higher education, foreign language teaching methodology, translation studies, and learner autonomy development.

Phan Thi Thanh Tam holds a Master's degree and is currently a lecturer at the University of Labour and Social Affairs, Viet Nam. She has over 25 years of experience in teaching and academic work. Throughout her career, she has been actively involved in teaching English language courses and developing innovative teaching methodologies. She has also contributed to academic research with several scientific publications in reputable national journals. Her research interests mainly focus on English language studies, language teaching methodologies, and innovations in foreign language education.

Dung Thanh Nguyen is a teacher of English and dean of General Education Department at National University of Arts Education in Vietnam. She obtained a bachelor degree in TESOL at Vietnam National University (VNU) in 2001, a master degree in TESOL at VNU in 2009, and a doctor degree in linguistics at Vietnam Academy of Social Sciences in 2017. She has published articles about English teaching issues such as implementing new methods in classrooms and ways of motivating EFL students. She is interested in English teaching methodology, motivation, learning styles and materials and curriculum development.

Uyen Thu Thi Duong is currently a second-year student majoring in English language, specializing in Economics and Business. She has actively participated in academic and extracurricular activities, including scientific research projects and writing reports for scientific conferences. She also has experience in public speaking and event hosting through volunteer and educational programs. Her academic interests mainly focus on English language studies, communication, economics, and business.

Appendices

Appendix A

Measurement items

Construct	Code	Measurement items
Engagement (ENG)	ENG1	I actively engage with English vocabulary content on TikTok and Instagram Reels.
	ENG2	Short-form video platforms increase my motivation to learn English vocabulary.
Repetition (REP)	REP1	I frequently encounter repeated English vocabulary in short-form videos.
	REP2	Repeated exposure helps me remember English vocabulary more effectively.
Multimedia Elements (MUL)	MUL1	Visual and audio elements help me understand vocabulary more easily.
	MUL2	Subtitles and illustrations support my vocabulary learning process.
	MUL3	Multimedia features enhance my ability to memorize new vocabulary.
Distraction (DIS)	DIS1	Entertainment-oriented content distracts me from vocabulary learning.
	DIS2	Rapid video transitions reduce my concentration while learning vocabulary.
Vocabulary Acquisition (VOC)	VOC1	Short-form videos help me expand my English vocabulary.
	VOC2	I can remember vocabulary more effectively through short-form videos.
	VOC3	My vocabulary usage has improved after using TikTok and Instagram Reels.

Scale:
1 = Strongly disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly agree

Appendix B

Demographic questionnaire

1. Gender

□ Male
□ Female

2. Age

□ Under 18
□ 18–22
□ Above 22

3. English proficiency level

□ Beginner
□ Intermediate
□ Advanced

4. English learning experience

□ Less than 1 year
□ 1–3 years
□ 4–6 years
□ More than 6 years

5. Frequency of self-learning English through short-form video platforms

□ Rarely
□ Sometimes
□ Often
□ Very often
□ Always

6. Which short-form video platform do you use most frequently for English learning?

□ TikTok
□ Instagram Reels
□ YouTube Shorts
□ Facebook Reels
□ Others: _____

7. Average daily time spent watching English-learning short-form videos

□ Less than 15 minutes
□ 15–30 minutes

□ 30–60 minutes
□ More than 60 minutes

#128 USABILITY TESTING OF WEB-BASED DICTATION TOOLS FOR ESL LISTENING COMPREHENSION

Thao Nguyen Thi Da[1], Ngoc Phung Tuan[2]

Abstract

Dictation has long been recognized as an effective method for improving ESL learners' listening comprehension. With the rise of web-based applications, learning English has become more accessible and engaging. However, the abundance of dictation websites can pose challenges for both teachers and students in selecting suitable platforms. This qualitative usability case study examines the affordances of three popular web-based applications – DailyDictation, English Club, and Parroto – through usability testing focused on site design and user experience. Ten non-English-major university students with mixed levels used the three websites over one week to complete usability testing tasks. Data was collected from users' journals, and visual analysis.

The study demonstrates that clear information and effective interaction design foster self-directed listening practice, enhanced by personalized suggestions, engaging activities, timely feedback, and error correction. These findings highlight the pedagogical importance of usability, urging educators and developers to prioritize adaptable, well-designed platforms that sustain long-term language learning outcomes.

***Keywords*:** dictation, web-based learning, listening, usability testing

1. Introduction

Listening is one of the fundamental skills besides speaking, reading, and writing. However, it is often overlooked as the ability to write and speak is predominantly viewed as the standard for language competency. In other words, listening is more like a minor skill which is a means to achieve speaking proficiency (Nunan, 2002). The perspective in which listening is static and passive has influenced the way this skill is taught. Ji (2015) reflected the traditional approach of listening teaching in many schools with a rigid procedure including the introduction of complex lexicon and background knowledge, followed by multiple listening exercises. As a result, students play a more submissive role in learning since what they do is to patiently listen and respond to the oral stimulus from the tape recorder (Movva et al., 2022). Meanwhile, Nunan (1998) opposed the idea of "listeners as tape recorders" and suggested that they could construct and interpret audio input by utilizing their schematic knowledge. Listening thus requires more efforts from the learners' part and has received much attention since the theory of comprehensible input gained its dominance (Nunan, 2002).

Although listening emerged as crucial in language classrooms, acting as an input provider for students, promoting the language skill among non-majors in universities and colleges is not a

[1] *Master, English Division, Faculty of Fundamental Sciences, Sai Gon Technology University, Ho Chi Minh City, VietNam; Email: thao.nguyenthida@stu.edu.vn*

[2] *Master, English Division, Faculty of Fundamental Sciences, Sai Gon Technology University, Ho Chi Minh City, VietNam*

straightforward task (Hong et al., 2024). Students are most often hindered in listening comprehension due to a dearth of lexical, grammatical, and schematic resources. The voices and accents of speakers also remarkably contribute to the grasp of the discussed notions if they present a wide array of phonological features. Other affecting factors to be considered include anxiety, non-verbal clues, and listening contexts (Ji, 2015; Movva et al., 2022).

To address these challenges, the present study revisits dictation – a conventional listening practice – through its integration with three web-based applications: DailyDictation, English Club, and Parroto. Unlike prior research that has primarily examined the effectiveness of digital tools in improving listening comprehension, this study advances the discussion by focusing on usability. Specifically, it investigates how accessibility, presentation, task diversity, and individual differences shape learners' willingness to engage with dictation websites. In doing so, the study evaluates the usability and suitability of these platforms for non-English-major students, aligning its scope with the following research questions:

1. What features of web-based applications do non-English-major students prefer regarding dictation activities?

2. Which of the three web-based applications – DailyDictation, English Club, and Parroto – is the most suitable for non-English-major students?

2. Literature review

2.1 The current trend in teaching Listening

Students in many language classrooms are of diverse abilities and levels, and while there is no one-size-fit-all way to deliver listening lessons (Gower et al., 2005), it is vital to tackle such differences with the most current teaching methods. A well-known approach is Activity-based language teaching, which comprises pre-listening, while-listening, and post-listening activities (Movva et al., 2022). This three-stage format is inherently based on the traditional model characterized in mainstream language teaching manuals such as Richards & Renandya (2002), Field (2009), or Parrish (2019). Although the stages persist throughout the development of language teaching, it is the core activities that have evolved, from the fundamental steps recognized by Mary Underwood (including the activation of learners' vocabulary and background knowledge, the development of information retrieval, and the consolidation of comprehension) to the adoption of task-based listening activities that emphasized the information transformation in real-world communication (Cook, 2013). For instance, Parrish (2019) suggested using a discussion or questionnaire to generate learners' schemata in addition to pre-teaching new words. However, not all researchers are in favor of the conventional activities with pre-set questions. Ji (2015) argued that this approach significantly hinders listening improvement and learners' motivation as it emphasizes the outcomes instead of the process. This author adopted strategies from the realm of discourse analysis to solve learners' difficulties in the listening process, such as anxiety, speech presentation, or message gaps. In the same vein, He (2016) modified Underwood's model using the "lexical chunks schema-oriented listening teaching method" to enhance short-term memory and bridge learners' background knowledge with the listening materials.

While scholars hold different perspectives in terms of techniques in teaching listening, they seem to agree on the use of authentic materials (Mendelsohn, 1998; Richards & Renandya, 2002; Field, 2009). According to Ur (2024), the ultimate goal of teaching listening is to prepare students

for real-life situations where speech is generated spontaneously, featuring background noises, multiple accents, and redundancy, rather than formal pre-scripted discourse followed by a series of comprehension tasks. To make listening lessons more authentic, the author suggests two approaches. One involves utilizing the teacher's talk as a valuable source of language input that can scaffold learners to accomplish language tasks or offer opportunities for language acquisition (Moser et al., 2011; Jin & Webb, 2020). The other necessitates the proper choice of teaching materials, whether they are recordings from official coursebooks or online video tapes. Depending on the purposes, teachers can choose to make the listening texts or tasks authentic. Nunan (2015) believed modifying the task is even more crucial than editing the listening text, and thus, there is room for either authentic or non-authentic materials. Evidently, Chang et al. (2018) showed students' listening improvement using audio files from graded readers in combination with extensive listening and reading-while-listening techniques. The difficulty degrees of listening activities were also recognized in Hong et al. (2024) to cater to obstacles that learners often encounter, including a limited lexical resource, poor sound recognition and distinction, lack of listening strategies, and low concentration.

2.2 Dictation and online listening practice

In addition to the prevalent three-stage approach to listening teaching, dictation has emerged as a promising technique to enhance this language skill. The origin of dictation can date back to the Middle Ages when it was strictly associated with the grammar translation method in foreign language teaching. The technique underwent ups and downs over the nineteenth century. As the natural method became fashionable, it promoted more imitation in speaking and listening and, thus, dictation was rejected. However, it regained its high status with the introduction of the direct method which featured phonetic dictation exercises, before being overshadowed by the remarkable success of the audio-lingual method later (Stansfield, 1985). Nation & Newton (2009) argued that despite its peaks and troughs, dictation can make tremendous contribution to the language proficiency, particularly in vocabulary learning. Dictation urges learners to focus more on forms and structures at the clause and phrase levels by allowing them to be aware of their errors through checking the dictation on their own, in pairs or groups. When choosing dictation materials, it is important for teachers to take accent variety into account. This is because English has been widely used by native speakers of other languages, making it a lingua franca with various norms that are different from the standardized rules (Kecskes, 2019). Nowadays, deploying listening resources from English native speakers is not so difficult due to the omnipresence of multimedia instructional applications. Najjar (1996) depicted multimedia is the utilization of "text, graphics, animation, pictures, video and sound to present information". Therefore, these means can be easily accessed through a computer or other portable devices such as smartphones or laptops. Media-rich classrooms may facilitate a more engaging and motivating learning experiences as media-language learning tools are able to bring the real-life world into the conventional classroom, complementing it with interactive interface accompanied by sound, animation and other visual and aural effects (Mukhtarkhanova et al., 2023). Such multimedia, additionally, can be modified to match leaners' needs and preferences (R. Liu, 2023). For instance, visual learners may benefit most from watching videos or graphs, while auditory learners can make use of podcasts.

Online learning is one of the ways to access these multimedia tools, offering more exposure to effective L2 listening. First of all, students are able to work in their own space by controlling the listening speed, repeating the audio, or reading the caption (Robin, 2007). They can also access

an inexhaustible collection of listening activities, topics and accents (Kamalitdinovna, 2021). Therefore, the impact of online listening is significant in terms of learning strategies and incidental vocabulary acquisition, respectively (Smidt & Hegelheimer, 2004), concurs with Rahmat (2018) who discovered that complementing multimedia learning with dictation can boost students listening comprehension. However, the author warned that students in the experimental group encountered difficulty at the beginning of the treatment, saying that they cannot comprehend the recorded native voice of speakers. This somehow reflects the findings in Kazazoğlu (2013), which indicated a more excessive number of word errors from tape-recorded dictation compared to teacher-led one. Furthermore, if metacognitive instructions are integrated in the online listening tasks, the L2 listening comprehension ability can be improved as learners are encouraged to use strategies such as evaluation which can be achieved through dictation (Pei et al., 2023).

2.3 Framework for assessing a web-based application

When it comes to using web-based applications to assist language learning, there are a variety of choices online. Some have been proven to effectively improve learners' language skills, such as Wordwall, Memrise for boosting vocabulary acquisition (Esmaeili & Shahrokhi, 2020; Hasram et al., 2021; A. T. Nguyen et al., 2023), LyricsTraining for listening enhancement (Borromeo García, 2015; T. G. M. Nguyen, 2023), ELSA for communication practice (Dhivya et al., 2023), or Duolingo for reading and listening proficiency (Jiang et al., 2024). Such an abundance of digital resources, while adding diversity and interest to language lessons, can cause decision fatigue and potentially put considerable pressure on teachers. Web-based tools, in a strict view, can be considered teaching materials, and the decision on the choice of which to use is completely subjective, and the procedure for evaluating an adequate material requires an exhaustive list featuring principles from theories of language teaching and learning, and language acquisition (Tomlinson, 2023). However, learners' experience with digital materials is rather different from the traditional ones, demanding another perspective for evaluation.

Erdogan & Uzun (2010), while claiming there is an obvious neglect of pedagogical and methodological approaches, developed a checklist for an adequate language learning website. The evaluation checklist features general information about the website, user interfaces, contextual aids, and educational purposes. Meanwhile, M. Liu et al. (2015) proposed usability testing as the framework for assessing a website's design, which utilizes a series of interaction tasks for users to explore the website, experience a specific feature, and navigate through different sections. Moreno Fuentes & Risueño Martínez (2018) suggested using a checklist from the perspective of ergonomics, linguistics, and pedagogy. The checklist includes nine areas: general site information, language skills and components, educational material, multimedia use, interactivity, communication, aid tools and linguistic resources. In the same vein, Alhabdan (2021) took the viewpoint of students and developed a checklist questionnaire with two categories: functionality and usability. The former refers to the ease of access and navigation, while the latter involves the design, content, and instructions offered by the website.

Overall, to evaluate a language learning website, previous research often targets the user interface and the quality of online materials. Hadjerrouit (2010) suggested that any web-based learning resource should include four important features:

1. It is hosted online through websites

2. Its content serves specific learning outcomes
3. Its design involves theories from the learning and teaching fields
4. It is reusable for various contexts.

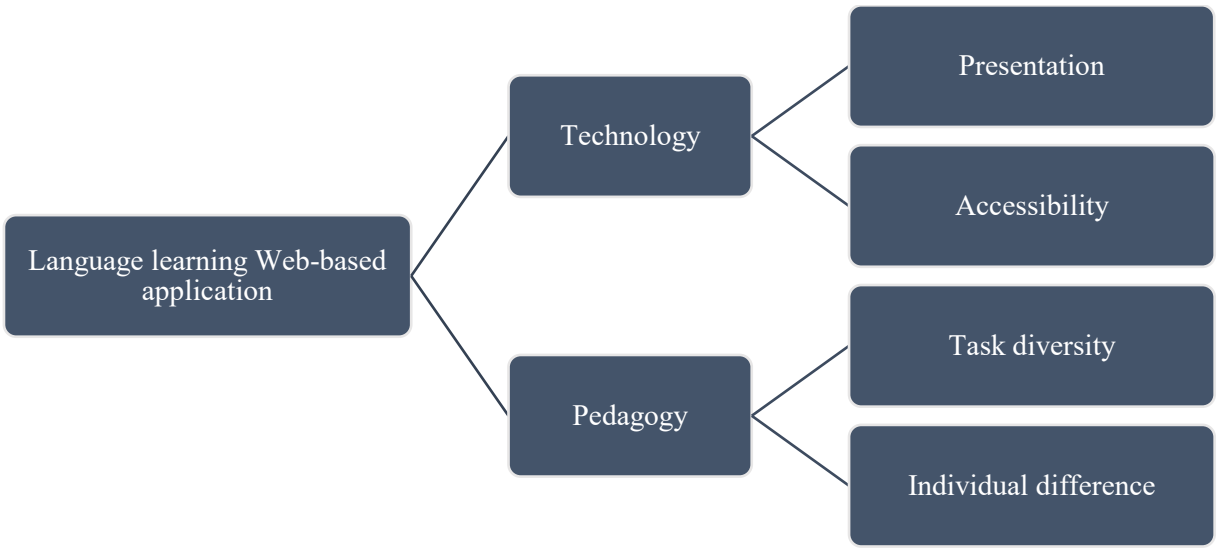
The author stressed that online learning tools were still in the ground of software programmers instead of teachers or educators, and therefore, their "technical usability" seems more dominant and influential than "pedagogical usability".

Although there are concerns about linguistic and pedagogical values, the link of those elements to learning and teaching theory seems to be overlooked. For example, the characteristic of helping learners to conduct language tasks step by step and facilitating interaction with other people is well aligned with Vygotsky's theory of mediation and zone of proximal development (Hadjerrouit, 2010; G.-Z. Liu et al., 2011; Huong & Bui, 2024). Moreover, in order to help learners to be more independent and self-regulated beyond the classroom, they should fully know their strengths, weaknesses, preferences, and styles to make informed learning decisions (Hadjerrouit, 2010; Brown & Lee, 2015; Allison et al., 2019). These individual differences are believed to be closely connected to the domain of second language acquisition. Apart from these concerns, some other aspects need attention as well. Web users nowadays can access online platforms through various devices such as smartphones, tablets, or laptops with different screen sizes, demanding the ability to optimize the interface for effective interaction. Furthermore, with the advent of AI technology, the integration of AI models into existing learning platforms should also be taken into account as it offers personalized learning experiences, allowing specific individual needs to be fulfilled through automatic assessment, real-time feedback, and educational recommendations (Shafiee Rad, 2025).

Based on the previous literature, a language learning web-based application should essentially balance technology and pedagogy (Figure 1).

Figure 1

The framework for language learning web-based applications



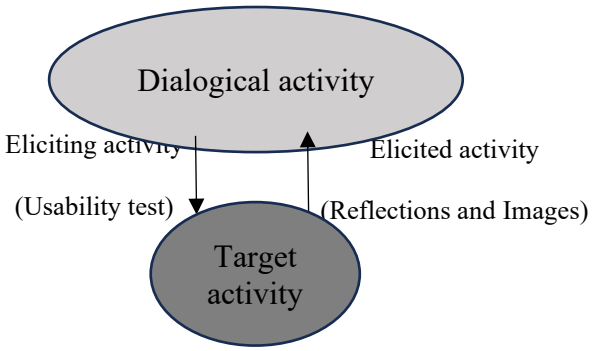
3. Methodology

This study adopted a small-scale qualitative usability case study to explore learners' experiences with three web-based dictation tools. Ten freshmen from Sai Gon University of Technology participated, all non-English-major students with mixed proficiency levels. They were selected through convenience sampling, as they were readily accessible to the researcher. The sample pool was limited to ten as the scope of this study is restricted to a case study and phenomenology (Creswell & Creswell, 2018). The students were asked to experience three popular web-based applications: DailyDictation, English Club, and Parroto. Meanwhile, they had to complete usability testing tasks and keep a journal which delineates their views and reflections on the applications. First, the authors met ten students in person, using a designed journal (Appendix) to instruct students how to reflect and take screenshots while experiencing the three websites. Then, a Zalo group was created for further communication between authors and participants. After receiving participants' journals via Zalo group, researchers would examine thoroughly and ask for clarification or changes if needed.

The data was collected mainly from participants' journals and photo-observation. In order to ensure students' reports adhere to the purposes and scopes of this study, a usability test was used to elicit their reflections for each application. The test was adapted from M. Liu et al. (2015), and it includes three categories: exploratory, specific, and open-ended tasks. The first type is for the participants to familiarize themselves with the applications, the second type requires them to conduct an assignment using some features of the applications, and the last one only asks them to perform a task without instructions. The use of the usability test to capture students' reflections was based on the eliciting model by Van Belleghem (2016) (Figure 2). The eliciting activity was the test, while the elicited activity was the constructed data – reflections written and photos captured in journals. The target activity was what the researchers expected the students to do in this study: recording their interactions with the applications, and the dialogical activity involves the member checking process where the researchers gave the participants the final report to determine whether they feel it was accurate (Creswell & Creswell, 2018).

Figure 2

Belleghem's model for reflection elicitation.



The data was analyzed using directed content analysis which can be applied to various writing materials. Texts are then summarized through the use of pre-determined categories, and thus, the theme structure usually emerges from the theories in the literature (Cohen et al., 2018). As mentioned in the previous section, there seems to be few coding schemes targeting learning and teaching fields, so this study filled the gap by synthesizing schemes (Table 1) from three sources in the literature (Hadjerrouit, 2010; Allison et al., 2019; Huong & Bui, 2024). While ten journal entries were read several times and coded in accordance with the pre-existing themes and subthemes related to the features of technology and pedagogy, the screenshot images were viewed as instances associated with the given themes.

Table 1.

Coding Scheme for the Usability Task

Coding Theme	**Explanation**	**Examples from Usability testing tasks**	**Examples from the current study's journal data**
Accessibility	The language learning tool is user-friendly and compatible with different devices; and the homepage indication, and navigation are supportive.	*Please take 5 minutes to explore the sites freely. If you need to register to explore, please do so.*	*The site runs smoothly without errors, the homepage is easy to view and use.*
Presentation	The content such as corlors, fonts, styles, media or graphics is easy to read and impressive.	*Pay attention to the interface design (labels, colours, and banners), ease of navigation, and the types of listening materials available.*	*Animated interface, cartoon videos appeal to GenZ, allowing users to learn while being entertained.*
Task diversity	The activities provided by the online application are varied including games, quizzes, riddles, pedagogical or real-life tasks.	*Use the site to find a dictation that matches your current English level.* *After completing it, explore the related 'Dictation Tips' or 'Grammar' links.* *Decide if these additional resources help you understand why you made certain mistakes.*	*Listening exercises, quizzes and games in each section*
Individual Difference	Learners' needs, characteristics and learning styles are taken into account to optimize learning experiences. AI technology is integrated to support learners' autonomy by providing timely and specific feedback for their actions.	*Pick a short video clip of your interest.* *Use the 'repeat' or 'slow down' features to transcribe the dialogue.* *Verify your accuracy using the built-in feedback system.* *Report if the interactive features made the difficult parts easier to understand.*	*Easy to transcribe and replay multiple times. Accuracy check is quick and helps beginners improve their listening faster.*

To guarantee reliability, the usability test was piloted with a subset of four students of the same population. It is noted that these respondents did not participate in the main study. Based on the feedback of the students in the pilot study, the usability test and the instructions for journals were refined to suit the purpose of the current research. Moreover, the researchers also conducted content analysis on the pilot data and cross-checked with each other to reach the intercoder agreement. Specifically, each researcher coded the texts using the predetermined themes in Table 1 independently. After that, the results were compared and discussed to find the common grounds between the coders until the consensus of the coding reached at least 80% of the time (Creswell & Creswell, 2018).

4. Results

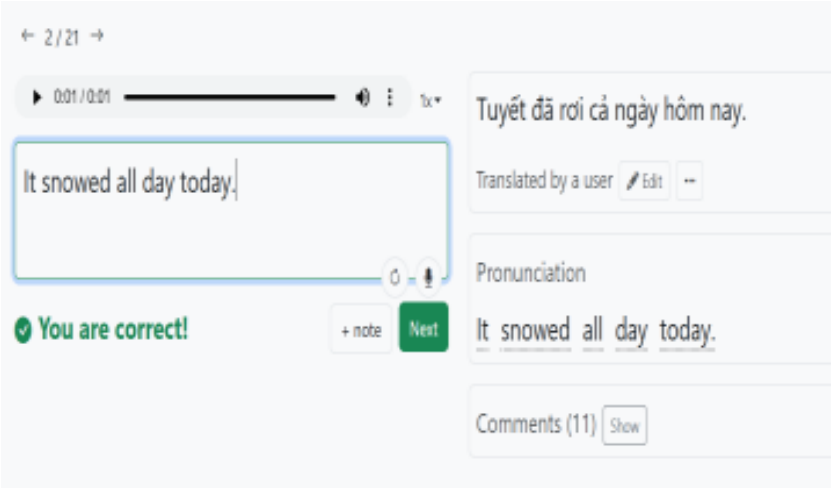
4.1 The Accessibility of the Online Listening Web-based Applications

The respondents in the current study have consistent perspectives toward the accessibility of the three platforms – DailyDictation, English Club, and Parotto. While the first and the last ones received positive feedback from the students, English Club did not seem to be user-friendly. Evidently, S1, S6, and S10 agreed that DailyDictation was "easy to use; accessible for newcomers." S3 and S8 shared a similar view about the "simple design" of this platform that makes it "easy to view and operate". This is also the first impression of S6 about it, she wrote, "First impression is simplicity, no flashy images, clear and easy to use. Feels serious for learning, not distracting." However, S1 reported that "the loading speed is slow and laggy; loading errors occasionally", which contradicts the experience of S2, S3, S7, S9 and S10 who observed "fast access speed", "smooth interaction", "normal speed without lag" and "quick loading time".

In the same vein, Parotto offered great usability as in the words of S2, "it is easy to view and search, with an online community chat bar for help if needed." S7 even stressed that this platform is "stable, easy to use, suitable for Gen Z and Alpha". Not all students have such positive experiences with this platform, S1 commented "There are serious errors. I cannot interact with the homepage to do exercises and cannot click on short clips to use". The loading speed was also a concern among the students. While only S5 wrote "the system runs stably. I'd give 9/10 for smoothness", the others, such as S3, S6, S7, and S9, described "the speed is a bit slow" or "the platform is sometimes laggy due to heavy traffic".

Figure 3

DailyDictation's feature of dictation activity



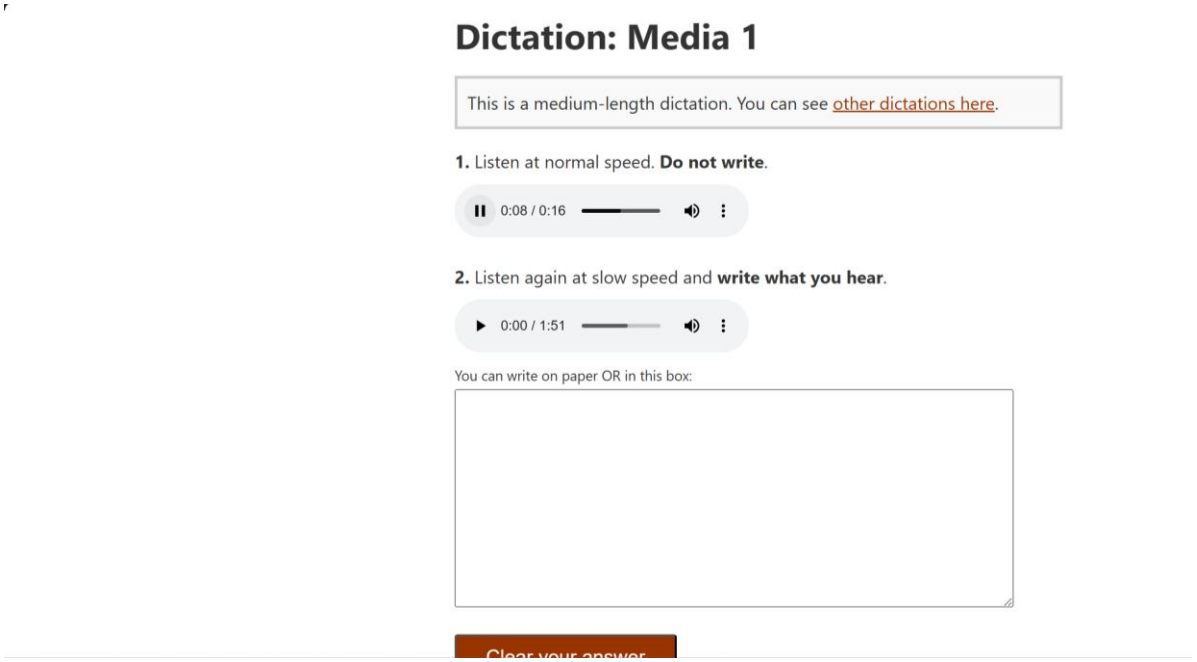
The participants appeared to be disappointed with the interaction with English Club when most of them reported undesirable experiences with the loading speed, search engine, and device compatibility. S2 wrote in her journal, "Although divided into clear sections but lacks illustrative images for beginners, the site speed is slow, and the entertainment value is low. It needs a cleaner interface for easier navigation. I do not want to return because it is hard to find the desired sections." S7 and S9 underwent opposite observations related to the compatibility elements. Whereas S7 noted, "the user interface fits phones only, and it is uncomfortable on desktops. It is hard to locate exercises, and the audio is unclear for full sentences", S9 wrote, "the application is not optimized for mobile. The experience is not smooth. It needs improvement." In the worst-case scenario, S1 reported "the anti-bot mechanism makes it difficult or impossible for users to access."

4.2 The Presentation of the Online Listening Web-based Applications

As for the presentation theme, DailyDictation and Parotto continued to be preferable. The former was viewed as an "interesting site" (S1) with "pleasant graphics, high entertainment value" (S2), "clear interface" (S3), "exercise organized by topics, clear readings" (S4). S7 and S10 acknowledged the presentation modes of the application; for example, S7 thought "the app has a harmonious, minimalist interface, and supports both light and dark modes". He also pointed out that "the 'Available exercises' section looks glaring with green font." Some other students valued the quality of audio and visual features, such as S5 and S9. S5 wrote "the site offers clear loud voices. No distortion, smooth use without lag, even with simultaneous traffic," and S9 mentioned "short listening clips prevent boredom". Furthermore, the display of scores and results was considered "easy to view" by S3 and "clear, with both Vietnamese and English for better understanding" by S5.

Figure 4

Englishclub's feature of dictation activity



Parotto significantly impressed the participants with its design as in the words of S5, "Tech-style design makes it engaging for users." S2, S6, and S8 also mentioned the "attractive interface", "clearly different, modern design", and "dynamic design" in their journals. Specifically, S2 stressed the entertainment value when she wrote "animated interface, cartoon videos appeal to Gen Z, allowing users to learn while being entertained". S6 and S8 emphasized the "cute, eye-catching design, and rich materials, stimulating interest". Similarly, S5 detailed the strengths of Parotto thoroughly, "the layout is nicely connected with green-black tech style colors, engaging for learners. The display of visuals and audio is vivid and lively; words and quizzes are fairly easy. Fast-moving words are arranged to look modern and engaging."

When it comes to English Club, the respondents had mixed opinions about this platform, but the overall view is negative. S9 and S10 recognized "the content is divided clearly, from reading to writing, speaking, grammar, vocabulary". In the same vein, S6 and S8 also appreciated the rich content, but they pointed out that this was also a drawback. S6 complained that "it requires time to explore. Presentation is not very attractive, but it is still a great app for practice. There is lots of content, but the cluttered presentation makes it easy to lose focus". The complaints about the presentation of English Club were also present in the other students' journals. For instance, S3 commented that "the layout is cluttered with too much text, and the content is scattered and cluttered. It is hard to find listening exercises. The interface is hard to use, and the pronunciation is difficult to hear. The layout is too messy. I cannot find where to type text. It is very hard to use, so I cannot find the needed content". S7 remarked that the web design made him "feel academically rigid in a negative way".

4.3 The Task Diversity of the Online Listening Web-based Applications

An upside of DailyDictation is its wide range of topics. For instance, S7 wrote, "The app has many topics for different levels," and S8 and S9 shared this opinion, further stating, "it is worth using daily to improve writing and listening skills". Furthermore, speaking skills could be enhanced when students are asked to pronounce parts of the recording. While S1 claimed "it is easy to shadow", S2 "dislikes the pronunciation task". The app is especially suitable for low-level learners as S4 suggested that he could "improve English skills through everyday stories that are easy to understand". The length of the recording was what S6 considered useful

as she wrote, "short listening audio (1–2s) keep it engaging". Despite the topic diversity, the tasks offered by the application seem to be uniform. S2 and S5 commented on this shortcoming:

"The content is easy to listen to and understand, but exercises are somewhat repetitive". (S2)

"The site does well in basic areas, needs more new exercise types" (S5)

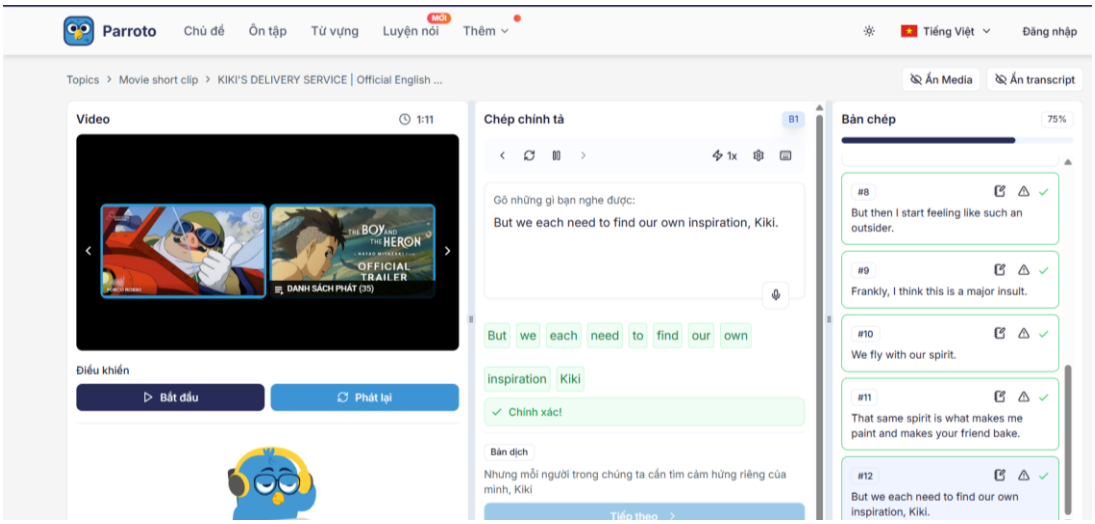
S8 suggested the application "should add features like short video transcripts".

The next issue with DailyDictation was the pronunciation task. Explaining why S2 found this section unfavourable, S3 wrote, "the pronunciation is good but quite fast, which may be difficult for weak English learners".

There was general agreement among the participants that Parotto is interesting due to its appealing tasks and exercises. As expressed by S8, "The website is convenient for dictation practice. It is more interesting than traditional dictation'". S3 also praised the website for being a game changer in dictation, "It is easy to find lots of interesting content, combining audios with videos/images makes dictation more engaging". S2, S4, and S7 shared similar opinions; for example, S4 reported his experiences with various task formats, "The site offers many interesting learning types such as short clips, music, games, and exam practice". The combination between entertainment and learning was another good aspect of Parroto. According to S6, she stated "Learning is fun through videos, practicing listening, dictation, and speaking. It includes pronunciation practice with native speakers. It makes dictation practice fun by combining audio, video, and images". S9, on the other hand, associated the application to "a mini social network, with vocabulary, speaking, dictation videos, grammar, mini games".

Figure 5

Parroto's feature of dictation activity



Task diversity is also a recurring theme in the students' journals for English Club. Most respondents indicated that the website offered quizzes, games, and videos to make learning less boring. One student (S9) explained that "There are mini games, pronunciation videos, and supports for TOEIC test preparation. It explains new words, and includes interactive exercises, and diverse resources". Another (S4) noted the effectiveness of the application for English skills improvement, "It helps users improve many English skills such as listening, speaking, reading, and writing. It offers many short exercises that help users practice daily without

boredom". In contrast, though S7 agreed that English Club featured "many fun activities like games and quizzes, its content was outdated".

4.4 The Individual Differences of the Online Listening Web-based Applications

There were slightly divergent opinions regarding the individual differences of DailyDictation. While most students recognized the versatility of the application, a few others complained about the audio recordings. In terms of flexibility, S1 shared his view that, "It is very worthwhile, and recommended for self-study or shadowing. The app is easy for beginners to access". A similar response was also observed in S2 and S6's journal, which said "It is easy to correct mistakes and remember, especially suitable for weak English learners" and "The application is suitable for those with basic listening skills who want to upgrade to higher levels". S7 and S8, on the other hand, believed that DailyDictation included "multiple listening modes suitable for most users". This was explained in the S4 report, "There are many different topics, serving various learning purposes, age groups, and exam preparation needs. What I like most is the answer check feature after listening, and the voice chat function that helps users conveniently improve their speaking skills. There are answers for learners to correct mistakes themselves". However, there were issues related to audio recordings that the application developers need to address. S3 commented that "the speed adjustment mode makes it usable, and easy to access and use, but it is not very engaging, may get boring quickly". Another respondent who also shared this viewpoint further stated that she did not want to return to this website because the speaker swallows words in some recordings. It should provide clearer voices for beginners, and reduce excessive hints in answer checking". Regarding the answer checking, this student suggested "it needs improvement because just one correct word reveals the whole sentence; hints should be reduced so users can replay and find answers themselves". Apart from the two main abovementioned aspects, other minor features contributing to the learner's difference were mentioned in the journals included:

"The app is entirely in English, difficult for beginners. It should add Vietnamese for easier understanding of new words." (S9)

"It includes interesting topics with clear audio. I can catch words after about 5 listening times. A good site that is suitable for 15-minute daily practice."

It was commonly reported that Parotto was a great application for students of any level and interest. As 3 noted, "Listening/pronunciation topics are interesting and match personal preferences. Lots of content matches personal interests, but some short videos require purchase. It is easy to transcribe and replay multiple times. The accuracy check is quick and helps beginners improve their listening faster. The application can be used for dictation practice and shared with others due to its engaging nature". Once again, the entertaining value of this application was mentioned in S2 and S4's journals. According to them, it is an "excellent website suitable for all ages, easy to learn, easy to understand, and highly entertaining", and "learning through images and videos is truly effective for me. It makes studying less boring and more engaging". Such positive responses can be explained by the report of S7:

"The application offers a good experience, diverse content, and it is easy to find interesting videos. Good interaction supports listening and knowledge. The application highlights mistakes and allows dictionary access."

In addition to the utility of Parotto, one notable emerging theme was collaboration, which allowed participants to connect and interact with their peers. For instance, S10 emphasized that "chat feature adds interaction, makes learning fun"; meanwhile, S8 appreciated the "friendly online chat community". These students also stressed the correction

feature of the application which "checks the correctness word by word, so there is no need to write full sentences".

English Club was able to cater to different learners' needs despite potential limits. As expressed by S5, the application was more appropriate for learners at higher levels. She wrote in her journal, "I like the answer-checking feature. The application was suitable for those with existing listening skills who want to improve further". Similarly, S8 stressed the lack of explanation for learners' errors when stating that "the application does not help understand mistakes. There is no error correction or pronunciation, only new vocabulary". Compared to the other websites, English Club also offered collaboration and various materials to serve different students' needs. This was evident in S4 and S7's comments:

"The application features community support."

"Within each skill, it provides clear definitions, improvement methods, effective learning strategies, and exercises for practice. It also offers links to YouTube, online eBooks, and reliable knowledge sources for further learning."

5. Discussion

The findings of this study demonstrate that the usability of web-based dictation tools is intrinsically linked to their pedagogical effectiveness. Although all three platforms aimed to improve listening comprehension, their success depended on the balance between *technical usability* and *pedagogical usability* (Hadjerrouit, 2010). Parroto emerged as particularly effective due to its technology-oriented design and multimedia integration, which aligns with the principles of Integrative CALL, where multimedia facilitates authentic practice and learner engagement (Chen & Zhang, 2011). This outcome resonates with prior research emphasizing the importance of authentic, interactive listening activities (Chang et al., 2018; Ur, 2024). Learners' appreciation of replay and feedback features further supports findings in MALL scholarship, which highlight autonomy and personalized support as key drivers of language learning (Huong & Bui, 2024; Lee & Xiong, 2023).

Conversely, English Club's cluttered and academically rigid interface was poorly received, echoing concerns raised by Erdogan & Uzun (2010) and Alhabdan (2021) regarding the neglect of usability in language learning websites. Despite its rich content, the platform's information overload contributed to frustration and decision fatigue, consistent with Lim's (2014) observations on the negative impact of excessive input. DailyDictation, while accessible for beginners, was criticized for its repetitive nature, which risks diminishing learner motivation. This finding reinforces Rost's (2000) argument that task diversity is essential for sustaining engagement. Taken together, the results suggest that effective dictation tools must not only function reliably but also provide multimodal, timely feedback that scaffolds learners within their Zone of Proximal Development (ZPD) (G.-Z. Liu et al., 2011; Huong & Bui, 2024).

For learners, the study highlights the importance of selecting platforms that balance technical and pedagogical usability. Features such as adjustable audio speed and instant accuracy checks empower learners to manage their own progress, reflecting the personalized feedback benefits noted by Golonka et al. (2014). More broadly, the findings underscore the growing significance of learner autonomy and self-regulation in digital learning environments (Alzieni, 2020; Brandt, 2020). Web-based dictation tools extend learning beyond the classroom, enabling students to tailor practice to their individual strengths, weaknesses, and learning styles.

For ESL teachers, the abundance of digital resources presents both opportunities and challenges. To avoid decision fatigue, teachers should adopt structured frameworks for usability evaluation rather than relying on subjective judgments. Applications that provide scaffolded activities (pre-, while-, and post-listening) and authentic input are more effective in preparing learners for spontaneous, real-world communication. Additionally, teachers should consider generational differences in technological preferences, as younger learners respond more positively to animated interfaces and vivid visuals than to text-heavy layouts.

For developers, the findings emphasize the pedagogical value of integrating adjustable playback, animated visuals, and timely feedback. Dictation tools should evolve beyond repetitive exercises to incorporate varied, authentic tasks that foster listening strategies and communicative competence. Ensuring device compatibility and seamless technical performance is also critical to maximizing pedagogical value.

6. Limitations

Several limitations should be acknowledged. First, the sample size was small, consisting of only ten non-English-major students in one university, which restricts the generalizability of the findings. Second, the study duration was relatively short, with participants using the platforms for only one week. This may not have been sufficient to observe long-term improvements or to determine whether the novelty effect of engaging platforms would diminish over time. Third, the reliance on self-reported journals and photo observations introduces subjectivity and potential bias. Finally, the study focused on only three specific web-based applications, meaning other potentially effective tools were excluded.

7. Recommendations

Future research should expand the sample size and include learners from diverse educational and cultural backgrounds to enhance generalizability. Longer testing periods would allow for the examination of sustained impacts on listening comprehension and learner motivation. Employing mixed-methods approaches that combine qualitative reflections with quantitative performance data could provide a more comprehensive evaluation of platform effectiveness. Researchers should also explore the potential of AI-driven personalization and adaptive feedback to optimize dictation tools for individual learner needs. For developers, recommendations include integrating varied and authentic tasks, enhancing device compatibility, and ensuring smooth technical performance to maximize pedagogical value.

8. Conclusion

This study examined the usability of three web-based dictation platforms – DailyDictation, Parroto, and English Club – through the dimensions of accessibility, presentation, task diversity, and learner differences. The findings underscore that well-designed dictation tools can meaningfully support ESL listening comprehension by offering clear interfaces, engaging activities, and responsive feedback. At the same time, technical shortcomings and repetitive exercises highlight the need for continuous refinement.

Dictation, long recognized as a versatile pedagogical technique, acquires renewed prevalence in the digital age when integrated into CALL and MALL environments. The results confirm its potential to foster attentive listening, vocabulary acquisition, and learner autonomy,

while aligning with broader literature on authentic input and interactive design. Importantly, the study demonstrates that usability is not merely a technical consideration but a pedagogical one: learners' ability to navigate, engage, and personalize their experience directly shapes their listening outcomes.

For educators, the implication is clear – dictation platforms should be selected not only for their content but also for their design quality and adaptability to learners' needs. For developers, the challenge lies in balancing technical innovation with pedagogical usability, ensuring that tools remain accessible, engaging, and responsive to diverse learner populations. Future research should extend beyond short-term usability testing to investigate long-term impacts, larger participant groups, and the integration of AI-driven personalization.

Ultimately, this study reinforces the enduring value of dictation as a bridge between listening comprehension and broader language development, while highlighting the importance of thoughtful design in maximizing the pedagogical potential of web-based learning tools.

References

Alhabdan, A. (2021). Evaluating English learning websites and digital resources from the perspective of Saudi English Language Learners: Technical Assessment. *Theory and Practice in Language Studies*, *11*(6), 639–651. https://doi.org/10.17507/tpls.1106.07

Allison, R., Hayes, C., McNulty, C. A. M., & Young, V. (2019). A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb. *JMIR Formative Research*, *3*(4), e14372. https://doi.org/10.2196/14372

Alzieni, H. (2020). The Impact of Mobile-Assisted Language Learning (MALL) in Developing the Listening Skill: A Case of Students at Dubai Men's College, the United Arab Emirates. *SSRN Electronic Journal*. https://doi.org//10.24093/awej/MEC2.6

Borromeo García, C. (2015). The Use of LyricsTraining Website to Improve Listening Comprehension. *MEXTESOL Journal*, *39*, 1–13.

Brandt, W. C. (2020). Measuring student success skills: A review of the literature on self-direction. In *National Center for the Improvement of Educational Assessment*.

Brown, H. D., & Lee, H. (2015). *Teaching by principles*. Pearson Higher Education & Professional Group.

Chang, A., Millett, S., & Renandya, W. A. (2018). Developing listening fluency through supported extensive listening practice. *RELC Journal*, *50*(3), 422–438. https://doi.org/10.1177/0033688217751468

Chen, L., & Zhang, R. (2011). Web-Based CALL to Listening Comprehension. In *Current Issues in Education* (Vol. 13, Issue 4). http://cie.asu.edu/

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge.

Cook, V. (2013). *Second language learning and language teaching*. Routledge.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches.* SAGE Publications.

Dhivya, D. S., Hariharasudan, A., Ragmoun, W., & Alfalih, A. A. (2023). ELSA as an Education 4.0 Tool for Learning Business English Communication. *Sustainability*, *15*(4), 3809. https://doi.org/10.3390/su15043809

Erdogan, K., & Uzun, L. (2010). The Internet, language learning, and international dialogue: Constructing online foreign language learning websites. *Turkish Online Journal of Distance Education*, *11*, 90–107.

Esmaeili, Z., & Shahrokhi, M. (2020). The Impact of Memrise Application on Iranian EFL Learners' Collocation Learning and Retention. *International Journal of Language Education*, 222–233. https://doi.org/10.26858/ijole.v4i2.10672

Field, J. (2009). Listening then and now. In *Listening in the Language Classroom* (pp. 13–25). Cambridge University Press. https://doi.org/10.1017/CBO9780511575945.003

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, *27*(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

Gower, R., Phillips, D., & Walters, S. (2005). *Teaching practice handbook*. Macmillan Education.

Hadjerrouit, S. (2010). A Conceptual Framework for Using and Evaluating Web-Based Learning Resources in School Education. *Journal of Information Technology Education: Research*, *9*, 053–079. https://doi.org/10.28945/1106

Hasram, S., Nasir, M. K. M., Mohamad, M., Daud, Md. Y., Rahman, M. J. A., & Mohammad, W. M. R. W. (2021). The Effects of WordWall Online Games (WOW) on English Language Vocabulary Learning Among Year 5 Pupils. *Theory and Practice in Language Studies*, *11*(9), 1059–1066. https://doi.org/10.17507/tpls.1109.11

He, X. (2016). Non-English Majors' Listening Teaching based on Lexical Chunks Theory and Schema Theory. *English Language Teaching*, *9*(2), 142. https://doi.org/10.5539/elt.v9n2p142

Hong, T. L. T., Nhi, H. N. Y., & Nhi, H. T. Y. (2024). Boosting Listening Skills in Non-English Major Freshmen. *International Journal of Asian Education*, *5*(2), 112–122. https://doi.org/10.46966/ijae.v5i2.390

Huong, L. P. H., & Bui, H. P. (2024). English Listening via Online Applications: Tool-mediated Language Learning by EFL University Students. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(2), 814–837. https://doi.org/10.70730/KIFE1132

Ji, Y. (2015). Discourse Analysis and Development of English Listening for Non-English Majors in China. *English Language Teaching*, *8*(2). https://doi.org/10.5539/elt.v8n2p134

Jiang, X., Peters, R., Plonsky, L., & Pajak, B. (2024). The Effectiveness of Duolingo English Courses in Developing Reading and Listening Proficiency. *CALICO Journal*, *41*(3), 249–272. https://doi.org/10.1558/cj.26704

Jin, Z., & Webb, S. (2020). Incidental Vocabulary Learning Through Listening to Teacher Talk. *The Modern Language Journal*, *104*(3), 550–566. https://doi.org/10.1111/modl.12661

Kamalitdinovna, D. N. (2021). Teaching listening skills with online digital tools. *Mental Enlightenment Scientific-Methodological Journal*, *1*(01), 30–39. https://mentaljournal-jspu.uz/index.php/mesmj/article/view/48

Kazazoğlu, S. (2013). Dictation as a Language Learning Tool. *Procedia - Social and Behavioral Sciences*, *70*, 1338–1346. https://doi.org/10.1016/j.sbspro.2013.01.195

Kecskes, I. (2019). The Nature of English as a Lingua Franca. In *English as a Lingua Franca* (pp. 15–44). Cambridge University Press. https://doi.org/10.1017/9781316217832.002

Lee, J.-C., & Xiong, L. (2023). Exploring learners' continuous usage decisions regarding mobile-assisted language learning applications: A social support theory perspective. *Education and Information Technologies*, *28*(12), 16743–16769. https://doi.org/10.1007/s10639-023-11884-5

Lim, L. (2014). Engaging student interpreters in vocabulary building: Web search with computer workbench. *ReCALL*, *26*(3), 355–373.

Liu, G.-Z., Liu, Z.-H., & Hwang, G.-J. (2011). Developing multi-dimensional evaluation criteria for English learning websites with university students and professors. *Computers & Education*, *56*(1), 65–79. https://doi.org/10.1016/j.compedu.2010.08.019

Liu, M., Abe, K., Cao, M., Liu, S., Ok, D. U., Park, J., Parrish, C. M., & Sardegna, V. G. (2015). An analysis of social network websites for language learning: Implications for teaching and learning English as a Second Language. *CALICO Journal*, *32*(1), 113–152. https://doi.org/10.1558/calico.v32i1.25963

Liu, R. (2023). Analysis of multimedia technology and mobile learning in English teaching in colleges and universities. *Nonlinear Engineering*, *12*(1). https://doi.org/10.1515/nleng-2022-0300

Mendelsohn, D. J. (1998). Teaching Listening. *Annual Review of Applied Linguistics*, *18*, 81–101. https://doi.org/10.1017/S0267190500003494

Moreno Fuentes, E., & Risueño Martínez, J. J. (2018). Design of a checklist for evaluating language learning websites. *Porta Linguarum: Revista Internacional de Didáctica de Las Lenguas Extranjeras*, *2018*(30), 23–41.

Moser, J., Harris, J., & Carle, J. (2011). Improving teacher talk through a task-based approach. *ELT Journal*, *66*(1), 81–88. https://doi.org/10.1093/elt/ccr016

Movva, S., Alapati, P. R., Veliventi, P., & G, M. (2022). The Effect of Pre, While, and Post Listening Activities on Developing EFL Students' Listening Skills. *Theory and Practice in Language Studies*, *12*(8), 1500–1507. https://doi.org/10.17507/tpls.1208.05

Mukhtarkhanova, A., Ustelimova, N., Yessengaliyeva, A., Iliyassova, A., & Tanesh, B. (2023). The Use of Multimedia in English Classes as a Means of Increasing Student Motivation. *World Journal of English Language*, *13*(9), 10. https://doi.org/10.5430/wjel.v13n9p10

Najjar, L. J. (1996). Multimedia information and learning. *Journal of Educational Multimedia and Hypermedia*.

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.

Nguyen, A. T., Nguyen, T. T., Le, T. T., Phuong, H. Y., Pham, T. T., Huynh, T. A. T., & Nguyen, H. T. (2023). Effects of Memrise on Vietnamese EFL Students' Vocabulary: A Case Study at a College in a Rural Area. *Electronic Journal of E-Learning*, *21*(5), 450–460. https://doi.org/10.34190/ejel.21.5.3066

Nguyen, T. G. M. (2023). The Impacts of Lyrics Training on Improving Listening Skills for ESL Students. *AsiaCALL Online Journal*, *14*(1), 18–29. https://doi.org/10.54855/acoj.231412

Nunan, D. (1998). *Language teaching methodology: a textbook for teachers*. Longman.

Nunan, D. (2002). Listening in Language Learning. In *Methodology in Language Teaching* (pp. 238–241). Cambridge University Press. https://doi.org/10.1017/CBO9780511667190.032

Nunan, D. (2015). *Teaching English to speakers of other languages: an introduction*. Routledge, Taylor & Francis Group.

Parrish, B. (2019). Developing Listening and Speaking Skills. In *Teaching Adult English Language Learners: A Practical Introduction* (pp. 101–140). Cambridge University Press. https://doi.org/10.1017/9781009024709.005

Pei, T., Suwanthep, J., & Lu, H. (2023). The effect of self-directed online metacognitive listening practice on Chinese EFL learners' listening ability, metacognition, and self-efficacy. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1285059

Rahmat, A. Y. (2018). The Effect of Using Educational Multimedia in Dictation on Students' Listening Comprehension at MA Darul Hikmah Pekanbaru. *J-SHMIC: Journal of English for Academic*, *5*(1), 1–19. https://doi.org/10.25299/jshmic.2018.vol5(1).1096

Richards, J. C., & Renandya, W. A. (2002). Teaching Listening. In *Methodology in Language Teaching* (pp. 235–237). Cambridge University Press. https://doi.org/10.1017/CBO9780511667190.031

Robin, R. (2007). Commentary: Learner-based listening and technological authenticity. *Language Learning & Technology*, *11*(1), 109–115. https://doi.org/10.64152/10125/44092

Rost, B. (2000). Interaction Analyzed in Traditional and Satellite-Delivered Extension Educational Presentations. *Journal of Extension*, *38*(1). https://doi.org/10.66752/1077-5315.7923

Shafiee Rad, H. (2025). Reinforcing L2 reading comprehension through artificial intelligence intervention: refining engagement to foster self-regulated learning. *Smart Learning Environments*, *12*(1), 23. https://doi.org/10.1186/s40561-025-00377-2

Smidt, E., & Hegelheimer, V. (2004). Effects of Online Academic Lectures on ESL Listening Comprehension, Incidental Vocabulary Acquisition, and Strategy Use. *Computer Assisted Language Learning*, *17*(5), 517–556. https://doi.org/10.1080/0958822042000319692

Stansfield, C. W. (1985). A History of Dictation in Foreign Language Teaching and Testing. *The Modern Language Journal*, *69*(2), 121–128. https://doi.org/10.1111/j.1540-4781.1985.tb01926.x

Tomlinson, B. (2023). *Developing materials for language teaching*. Bloomsbury Academic.

Ur, P. (2024). Teaching listening. In *A Course in English Language Teaching* (pp. 99–112). Cambridge University Press. https://doi.org/10.1017/9781009024518.008

Van Belleghem, L. (2016). Eliciting activity: a method of analysis at the service of discussion. *Le Travail Humain*, *Vol. 79*(3), 285–305. https://doi.org/10.3917/th.793.0285

Appendix

Usability testing tasks

Exploratory task

Please take 5 minutes to explore the sites freely. If you need to register to explore, please do so. Pay attention to the **interface design** (labels, colours, and banners), ease of navigation, and the types of listening materials available.

Specific tasks

Daily Dictation	English Club	Parroto
1. Select a 'Short Story' exercise. 2. Adjust the audio delivery speed to 0.8x. 3. Type directly into the listening activity box. 4. Click the check button to receive your score breakdown.	1. Choose a 'Medium' dictation. 2. Listen to the audio once without typing. 3. Use the second and third readings to complete the text. 4. Check your answers against the provided transcript and identify one new vocabulary word you learned.	1. Pick a short video clip. 2. Use the 'repeat' or 'slow down' features to transcribe the dialogue. 3. Verify your accuracy using the built-in feedback system. 4. Report if the interactive features made the difficult parts easier to understand.

Open-ended tasks

Daily Dictation	English Club	Parroto
1. Find a topic that genuinely interests you (e.g., news, stories). 2. Practice with it for 10 minutes using your own strategy—such as shadowing or multiple replays. 3. Reflect on whether this site feels like a tool you would use daily for self-study.	1. Use the site to find a dictation that matches your current English level. 2. After completing it, explore the related 'Dictation Tips' or 'Grammar' links. 3. Decide if these additional resources help you understand why you made certain mistakes.	1. Search for a video clip related to a hobby you have. 2. Practice dictation with it. 3. After finishing, explain how the combination of visual (video) and audio cues helped or hindered your listening comprehension compared to text-only dictation.

#129 VIETNAM'S ESL TRANSITION: ADAPTING STRUCTURED LITERACY IN MICRO-LEVEL MATERIAL DESIGN FOR PRIMARY CLASSROOMS

Pham Trang Thu[1], Phan Hoang Thuy Linh[2]

Abstract

As Vietnam transitions from English as a Foreign Language (EFL) to a Second Language (ESL), primary students often face systemic phonetic barriers – specifically coda deletion, monosyllabic processing, and tonal interference – rooted in L1 influence. Despite this strategic shift, current communicative textbooks seem to lack the micro-level decoding methodologies required for ESL proficiency. This study investigates the compatibility of foundational Structured Literacy (SL) routines, extracted from the HMH Into Reading curriculum, with the Vietnamese phonetic context. Employing qualitative document analysis, the research utilizes a mapping matrix to evaluate how four core routines dismantle L1-induced errors. The analysis demonstrates that "Tap and Count" and "Sound-by-Sound Blending" can effectively disrupt whole-syllable guessing and enforce released stops, directly mitigating coda deletion. Furthermore, "Continuous Blending" and "Echo Reading" can facilitate coarticulation and recalibrate sentence-level prosody to override ingrained Vietnamese tonal habits. Ultimately, the study proposes a tripartite Adaptation Framework for Scripted Instruction, providing explicit "I do – We do – You do" bilingual scaffolds for Teacher's Guides. This research contributes a micro-instructional lens to material design, offering an actionable, systemic approach to bridge the gap between phonemic awareness and oral fluency, thereby establishing a foundation for future experimental interventions in ESL-emerging classrooms.

***Keywords*:** structured literacy, cross-linguistic influence, phonological awareness, scripted instruction, material design.

1. Introduction

1.1. Policy and Instructional Context

Vietnam is pursuing a shift from English as a Foreign Language (EFL) toward English as a Second Language (ESL) in schools, established as a national objective in Conclusion No. 91-KL/TW (Politburo, 2024) and consistent with the 2018 General Education Curriculum for English (Ministry of Education and Training, 2018). The EFL-to-ESL transition here refers to this educational ambition rather than an established sociolinguistic condition; at the primary level it increases the need for oral-decoding foundations supporting intelligible communication.

1.2. Vietnamese-Related Phonological Barriers and Research Focus

This study focuses on three L1-induced phonological barriers—coda deletion, monosyllabic processing, and prosodic transfer—examined through cross-linguistic influence (CLI), which concerns how previously acquired patterns affect additional-language processing and production (Odlin, 1989; Jarvis & Pavlenko, 2008). Each was selected for its theoretical link to Vietnamese–English structural differences and its correspondence to an observable instructional demand within foundational literacy routines. Communicative instruction is

[1] *University of Languages and International Studies, Vietnam National University, Ha Noi; Email: thupham01.work@gmail.com*
[2] *University of Languages and International Studies, Vietnam National University, Ha Noi.*

central to Vietnam's ESL-oriented ambition but may not specify how teachers should address coda omission, undifferentiated word processing, or L1-shaped prosody, whereas Structured Literacy (SL) offers a relevant micro-instructional lens (Fallon & Katz, 2020). This study therefore examines four routines from *HMH Into Reading Grade 1 Instructional Routines*: Tap and Count the Sounds, Blending: Sound-by-Sound, Blending: Continuous, and Echo Reading (Houghton Mifflin Harcourt, 2017). Prior work has documented Vietnamese learners' difficulties with final consonants, phonemic awareness, and prosody but has not mapped foundational decoding routines onto these barriers or converted such analysis into bilingual scripts; addressing this gap is the study's contribution.

1.3. Research Purpose and Questions

This qualitative document analysis investigates the theoretical compatibility of the four selected SL routines with coda deletion, monosyllabic processing, and prosodic transfer among Vietnamese primary learners. It further develops a Tripartite Adaptation Framework for bilingual I do – We do – You do teacher guidance. Two research questions guide the study:

RQ1. Which foundational routines in the *HMH Into Reading Grade 1 Instructional Routines* show theoretical compatibility with the L1-induced phonological barriers of Vietnamese primary learners?

RQ2. How can a conceptual adaptation framework be proposed to translate these routines into bilingual scripted instruction for Vietnamese teacher's guides?

2. Literature Review

2.1. Cross-Linguistic Influence and Foundational English Decoding

Cross-linguistic influence (CLI) is the principal lens for examining how Vietnamese learners' existing knowledge may shape their early processing and production of English. Odlin (1989) defines transfer as influence arising from similarities and differences between a target and previously acquired languages, and Jarvis and Pavlenko (2008) extend it to perception, production, and processing. Learners may not simply fail to produce unfamiliar English features; they may reconstruct English through patterns functional in Vietnamese but insufficient for English decoding and oral fluency.

CLI should not, however, explain every non-target-like production: Selinker's (1972) interlanguage recognizes evolving approximations not all attributable to the first language, so the study focuses only on its three barriers. CLI also bears on early decoding—Koda (2005) argues that second-language reading builds on prior literacy experience, learners approaching a new writing system with established procedures—which suggests English phoneme-by-phoneme decoding may require explicit attention where familiar L1 routines do not foreground individual phonemes within a monosyllable.

2.2. Vietnamese–English Phonological Barriers

2.2.1. Coda deletion and final-consonant production

A frequently identified difficulty for Vietnamese learners concerns word-final consonants. Vietnamese permits a more restricted range of codas than English, so learners may omit, reduce, or reshape English word-final material. Nguyen (2007) identifies final consonants as a significant difficulty, and Disney and Le (2024), analysing speech from sixteen Vietnamese adult learners, found expected word-final /s/ and /z/ omitted in 28.4% of cases, particularly in morphologically marked forms and consonant clusters. Their findings show omission cannot be attributed solely to L1 transfer, since phonological complexity and

morphological function also matter. Because this evidence concerns sibilants and morphological markers rather than stops, its extension to final-stop instruction here is an analytical inference rather than a direct empirical claim.

Although these studies involve older learners, they provide a defensible basis for coda-focused instruction at the primary level. In beginning decoding, retaining a final consonant is necessary for producing the full sound sequence a word represents: omitting the final consonant in sit or cat loses part of the word form. The objective is not that every English final stop be audibly released in natural speech, but that early instruction make final stops salient enough to notice, retain, and practise as components of the target word—a distinction that matters when analysing routines for coda-related errors.

2.2.2. Monosyllabic processing and phonemic awareness

The second barrier concerns learners' access to the internal sound structure of short English words. Foundational literacy relies on phonemic awareness: learners must recognize that a word such as cat consists of separate phonemes that can be identified and blended. Learners entering through an established L1 system may not automatically employ the same phonemic decomposition, and Koda's (2005) approach supports the view that prior literacy experience influences additional-language word processing. Monosyllabic processing is therefore a plausible instructional challenge; learners may need explicit support moving from repeating a complete syllable toward identifying individual phonemes rather than a measured characteristic of all learners; it grounds the examination of whether materials make phoneme-level analysis observable and teachable.

This issue is closely related to coda deletion: where a short word is treated as an undivided oral form, a vulnerable final consonant may go unnoticed or be omitted without correction, whereas a routine requiring learners to account physically or orally for each phoneme makes the final sound part of the expected response. Phonemic segmentation is thus relevant to decoding accuracy and to more stable word-final production.

2.2.3. Prosodic transfer and connected oral reading

The Vietnamese–English contrast also operates at the prosodic level. Vietnamese uses syllable-level pitch contrastively, whereas English relies substantially on stress, prominence, rhythm, and intonation across syllables, words, and phrases. Because lexical tone and metrical stress are distinct systems without direct equivalence, the study uses the broader term prosodic transfer for the influence of L1 prosodic experience, including the salience of Vietnamese syllable-level pitch on learners' processing of English prominence and connected speech, rather than claiming lexical tone is transferred onto English stress.

Le (2021) found that L1 Vietnamese speakers differed from English-speaking participants under particular stress and intonation conditions, indicating that Vietnamese prosodic experience may influence learners' interpretation of English prominence cues. Nhat and Hien (2025) treated sentence stress, intonation, and connected-speech features as important dimensions of Vietnamese learners' pronunciation development. These studies do not establish how Vietnamese primary children read English sentences aloud, but they support including sentence-level oral modeling here. If prosodic transfer concerns the carry-over of L1-shaped prosodic habits into English, learners require clear English models that can be heard, repeated, and internalized; an oral-reading routine requiring imitation of modeled pronunciation and expression is therefore theoretically relevant to the prosodic dimension.

2.3. Structured Literacy and Scripted Scaffolding

Structured Literacy (SL) is relevant because it emphasizes explicit, systematic instruction in spoken- and written-language structures. Fallon and Katz (2020) describe SL as integrating spoken-language components, including phonology and sound–symbol relations into literacy teaching. Although discussed mainly in English-dominant contexts, its principles are pertinent here: learners are guided through sequenced procedures, features are made explicit, and feedback can target observable responses.

The study does not assume that an SL programme developed for native English-speaking children can be adopted unchanged in Vietnam, where primary learners acquire English as an additional language within a distinctive curricular and phonological context. The relevant question is whether particular foundational routines display mechanisms compatible with Vietnamese-related barriers; in the selected handbook, the four routines operationalize sound segmentation, sequential blending, continuous blending, and modeled oral reading (Houghton Mifflin Harcourt, 2017). The adaptation is also supported by the Gradual Release of Responsibility model, in which instruction moves from teacher modeling toward guided participation and independent performance (Pearson & Gallagher, 1983); the I do – We do – You do sequence provides the structure through which compatible HMH mechanisms become bilingual teacher-facing scripts.

2.4. Research Gap

Existing literature establishes that Vietnamese learners may have difficulty with English word-final consonants and prosodic processing, while cross-linguistic literacy theory supports attention to the transfer of prior routines into English decoding, and research on SL and gradual-release instruction grounds explicit, sequenced procedures. However, no located study systematically maps foundational SL routines onto Vietnamese phonological barriers and converts the compatibility judgments into bilingual scripted guidance for primary teachers. The study addresses this gap by analysing the four HMH routines against coda deletion, monosyllabic processing, and prosodic transfer. Its contribution is conceptual and material-oriented: a theoretically grounded basis for adapting explicit micro-instructional routines to Vietnam's ESL-emerging primary context and future classroom validation.

3. Methodology

3.1. Research Design

This study adopted a theory-informed qualitative document analysis to examine the compatibility of selected foundational Structured Literacy (SL) routines with L1-induced phonological barriers relevant to Vietnamese primary learners. Document analysis suits this aim because it enables systematic examination of stated purposes, procedural sequences, model language, and implied learner actions (Bowen, 2009), and it fits an exploratory, theory-building phase that necessarily precedes classroom intervention: a routine's theoretical compatibility with a barrier must be established before its classroom effects can be tested. The analysis was conceptual rather than experimental, examining whether observable mechanisms aligned with the framework's barriers rather than measuring learner performance. Compatibility was defined as the extent to which a routine required learners to perform a phonological or oral-reading operation relevant to the targeted difficulty; the focus was thus on instructional affordances—what each routine makes learners attend to, produce, combine, or imitate.

3.2. Documentary Corpus and Routine Selection

The primary corpus was the HMH Into Reading Grade 1 Instructional Routines handbook (Houghton Mifflin Harcourt, 2017), selected for its explicitly structured routines, stated purposes, materials, steps, and model teacher language, supporting micro-level analysis of how phonological instruction is operationalized. Four routines were purposively selected for their direct correspondence to the study's abstract and research questions: Tap and Count the Sounds, Blending: Sound-by-Sound, Blending: Continuous, and Echo Reading. A routine was included only when it explicitly involved phonological or oral-reading activity, contained observable procedures, and could be mapped onto at least one targeted barrier. Because the analytical interest lies in mechanisms rather than labels, the unit selected was the procedure exhibiting the relevant mechanism, which in one case is a step within a larger routine: the tap-and-count procedure is documented within the High-Frequency Words routine rather than as a discrete routine, and is treated here as a focal unit for its phoneme-segmentation mechanism (see Section 4.2, note). Other routines were excluded when their primary mechanism did not bear on the three barriers: the Vocabulary routine (p. R12) targets word meaning rather than phonological operations; the Syllabication: VCCV Pattern routine (p. R17), though it involves blending, operates on two-syllable words and so falls outside the study's monosyllabic scope; and Choral Reading (p. R25) was treated only as a possible supplement to Echo Reading, since the abstract specifies four core routines. This screening preceded scoring, so selection was governed by mechanism and barrier scope rather than by an expectation of high scores. The analysis nonetheless remains subject to some selection bias, since the four retained routines were identified in advance as phonologically relevant; this is acknowledged in Section 5.4.

3.3. Analytical Categories and Unit of Analysis

The unit of analysis was the micro-instructional routine episode, a bounded sequence in which the source specifies a teacher action, an expected learner response, and a target spoken-language feature. Each episode was examined for its purpose, procedural steps, sample language, learner demand, phonological level, and relevance to the Vietnamese context. Three a priori barrier categories were retained from the fixed framework: coda deletion and final-stop production (difficulty retaining or articulating English word-final consonants); monosyllabic processing (treating one-syllable words as undivided units rather than phoneme sequences); and prosodic transfer (carry-over of Vietnamese-shaped prosodic patterns into English connected oral reading). These categories were not generated from the HMH material; they served as the lens for evaluating compatibility. Four focal episodes were delimited, one per routine; where a procedure recurred across host routines, it was treated as a single unit defined by its mechanism rather than counted separately.

3.4. Analytical Procedure

The analysis proceeded through four stages. First, each routine's procedural components were extracted: instructional purpose, teacher and learner actions, model word or phrase, and operating level (phoneme, word, or connected oral reading). Second, each routine was coded for its primary mechanism and mapped onto the barrier it most directly addresses, with the routine-by-routine rationale presented in the Findings.

The four evaluation criteria were developed by the authors specifically for this analysis rather than adopted from an existing instrument. Their design follows a distinction common in materials evaluation, between a routine's intrinsic suitability for a learning objective and the practical demands of delivering it, separating two analytically independent questions: whether a routine's required learner action is theoretically aligned with a targeted barrier, and how

demanding it would be to adapt for a Vietnamese teacher's guide. Mechanistic alignment operationalizes the first, assessing whether the required action corresponds to the operation a barrier defines. Teacher feasibility and resource feasibility capture the practical dimension, the demand on the teacher's delivery and on materials, while adaptation requirement captures the modification needed to address L1–L2 distance. The 0–2 scale marks absent, partial, and strong presence of each criterion, avoiding a finer gradation that document evidence could not reliably support; the band thresholds (7–8 High, 4–6 Conditional, 0–3 Limited) let a routine be judged Conditional on practical grounds even when its mechanism was strongly aligned.

Third, each routine was scored on the four criteria using the 0–2 scale and band thresholds described above. Finally, routines assessed as compatible were grouped into phonemic segmentation, coda-to-word blending, and prosodic oral modeling, each operationalized through bilingual I do – We do – You do prompts informed by the Gradual Release of Responsibility model (Pearson & Gallagher, 1983; Fisher & Frey, 2014).

3.5. Trustworthiness and Scope of Interpretation

Analytical transparency was supported by an extraction-and-coding record linking every compatibility judgment to an observable routine feature; predetermined categories and explicit criteria reduced unsystematic interpretation. Coding proceeded in two stages: the primary author conducted the initial extraction and compatibility coding against the predetermined criteria, after which the second author reviewed each judgment against the corresponding documented feature. Discrepancies were resolved through discussion until agreement was reached, and the coding table was retained as an audit trail. Because the second author's review followed rather than preceded the initial coding, the procedure constitutes verification coding rather than fully independent double-coding; independent inter-coder analysis and expert review are identified in Section 5.4 as steps for future validation. Because the study analysed documents rather than outcomes, its findings concern theoretical compatibility and material-design potential. The active functions named in the abstract are therefore interpreted throughout as projected instructional functions requiring classroom validation, not demonstrated effects; where the abstract uses such verbs for concision, the findings qualify each as an instructional affordance evidenced in the documents. For the same reason, the abstract's term tonal interference denotes what this paper analyses more precisely as prosodic transfer (Section 2.2.3).

4. Findings

4.1. Overview of the Findings

The analysis identified a coherent relationship between the four routines and the three L1-induced barriers, monosyllabic processing, coda deletion and final-stop difficulty, and prosodic transfer, operating at different linguistic levels in a progression from phoneme awareness to word construction and connected oral reading. Consistent with the methodological scope, the judgments below describe projected instructional functions evidenced in the documents, not measured effects. As Table 1 shows, Tap and Count the Sounds and Blending: Continuous were rated High, whereas Blending: Sound-by-Sound and Echo Reading were rated Conditional because adaptation depends on additional teacher-facing support, not on any weakness in their mechanism.

Table 1

Compatibility of Selected HMH Routines with Vietnamese-Related Phonological Barriers

HMH routine	Documentary evidence	Targeted Vietnamese-related barrier	Identified mechanism	Judgment
Tap and Count the Sounds	Learners tap and count the sounds in can as /k/ /ă/ /n/ (p. R13).	Monosyllabic processing and whole-syllable guessing	Auditory-tactile phonemic segmentation	High, 8/8
Blending: Sound-by-Sound	Learners construct sit incrementally and include final /t/ before producing the whole word (p. R15).	Coda deletion and final-stop difficulty	Sequential phoneme accountability and coda-focused attention	Conditional, 6/8
Blending: Continuous	Learners blend mat without breaks before producing the complete word (p. R16).	Difficulty with word-level synthesis (inability to assemble identified phonemes into a complete word), with coda retention as a secondary concern	Continuous phonological synthesis	High, 7/8
Echo Reading	Learners repeat teacher-modeled oral reading while imitating pronunciation and expression (p. R24).	Prosodic transfer in connected oral reading	Modeled prosodic imitation	Conditional, 5/8

Note. Scores range 0–8 across four criteria. Bands: 7–8 = High, 4–6 = Conditional, 0–3 = Limited. Conditional ratings reflect implementation and adaptation demands, not weak mechanistic alignment; the mechanisms of Blending: Sound-by-Sound and Echo Reading are strongly aligned with their targeted barriers. Page references (R13–R24) are to Houghton Mifflin Harcourt (2017).

4.2. Tap and Count the Sounds: Making Phonemes Explicit within Monosyllabic Words

The first finding concerns the compatibility of Tap and Count the Sounds with the barrier of monosyllabic processing. In the HMH routine, learners are directed to represent each sound in the target word physically and orally. In the documented example, the word can is segmented into /k/ /ă/ /n/, with one finger-tapping movement assigned to each sound before learners identify the total number of sounds in the word (Houghton Mifflin Harcourt, 2017, p. R13)[1]. The analytical significance of this routine lies in the learner action it requires: a learner cannot complete the task by repeating can as a single memorized oral form, because the routine requires the word to be decomposed into three separately accountable units. This procedural demand corresponds directly to the study's definition of monosyllabic processing, in which a learner approaches an English one-syllable word as an undivided sound block rather than as a sequence of phonemes.

[1] *In the source handbook, "Tap and Count the Sounds" is not a standalone routine but a step within the High-Frequency Words routine (Houghton Mifflin Harcourt, 2017, p. R13; the worked example is can) and within Heart Words (p. R35). It is isolated here as a unit of analysis for its phoneme-segmentation mechanism; the label is retained for consistency with the abstract, and its application to general decoding is an analytical extension rather than a use prescribed by HMH.*

The routine scored 8/8 (High). Its mechanism aligned directly with the targeted barrier, the tapping action externalized phonemic structure observably, and it required no specialised resources and only minimal adaptation for bilingual guidance. Within the framework, Tap and Count the Sounds is therefore the clearest routine-level response to whole-syllable guessing, transforming internal sound analysis into a visible, monitorable classroom action.

4.3. Blending: Sound-by-Sound: Foregrounding the Word-Final Consonant

The second finding concerns Blending: Sound-by-Sound and its compatibility with coda deletion and final-stop difficulty. In the documented example, sit is not presented immediately as a complete oral form; learners are guided through a successive sequence in which individual sounds are introduced and blended until the final /t/ is incorporated (Houghton Mifflin Harcourt, 2017, p. R15). This makes the coda instructionally salient: the final consonant is neither optional nor embedded in undifferentiated whole-word repetition but is separately represented, elicited, and required in the completed blend. The routine thus activates sequential phoneme accountability and coda-focused attention, and a learner who omits the final /t/ produces an incomplete response the routine immediately exposes.

The mechanism of this routine is thus strongly aligned with coda deletion, consistent with the abstract's characterization of sound-by-sound blending as effective for foregrounding the coda. Its overall compatibility was rated Conditional (6/8) for implementation reasons rather than weak relevance: in the Vietnamese context, a teacher's guide would need explicit support for modeling final consonants clearly and avoiding productions in which the final sound is omitted or followed by an unintended vowel. The abstract's "released stops" should be read in an instructional rather than strict phonetic sense: the routine makes the word-final consonant salient and accountable within the decoding sequence, not audible release as a phonetic target, and does not establish how learners realize final stops across all connected-speech environments. Blending: Sound-by-Sound is therefore directly compatible with coda-focused instruction, provided implementation is supported by Vietnamese-specific teacher prompts and final-consonant monitoring notes.

4.4. Blending: Continuous: Moving from Segmented Sounds to Complete Word Production

The third finding concerns Blending: Continuous. In the HMH procedure, learners blend a word's sounds without breaks, as in the documented treatment of mat, before producing the complete word (Houghton Mifflin Harcourt, 2017, p. R16). Unlike the two preceding routines, it requires sustaining the full sound sequence in a single production. The mechanism, continuous phonological synthesis, is relevant to learners who identify individual sounds but fail to assemble them smoothly; it also supports coda-focused instruction when sequenced after sound-by-sound work, since a final consonant made explicit earlier must be retained within uninterrupted production.

The routine scored 7/8 (High): its mechanism was explicit, its resource demands low, and its core adaptable without redesign. The main adaptation concerned target-word selection—continuous blending should use words with vulnerable final consonants, so smooth production does not come at the expense of coda retention. Blending: Continuous thus links phoneme-level attention with complete word production, providing the transitional step between explicit coda attention and later oral fluency work.

4.5. Echo Reading: Providing a Model for Sentence-Level Prosody

The fourth finding concerns Echo Reading and prosodic transfer. The teacher reads a portion of text aloud and learners repeat it while imitating pronunciation and expression

(Houghton Mifflin Harcourt, 2017, p. R24). Unlike the earlier routines, it operates beyond isolated phonemes or words, targeting modeled oral production within phrases or sentences. The analysis identified modeled prosodic imitation as its central mechanism, corresponding to prosodic transfer because it provides an external English model against which learners' connected reading can be shaped—relevant to the abstract's claim that sentence-level prosody can be recalibrated through structured modeling and repetition.

Echo Reading received a Conditional judgment (5/8). It requires learners to imitate a model, but the source does not ensure the model suits Vietnamese learners' prosodic needs: its value depends on accurate teacher pronunciation, phrasing, prominence, and intonation, or on reliable recorded models and phrase-marked scripts. Echo Reading therefore carries the strongest implementation condition among the four, requiring more extensive adaptation than the phoneme- and word-level routines.

4.6. Integrated Finding: A Three-Level Instructional Progression

Across the four routines, the analysis identified a three-level progression: Tap and Count the Sounds supports phonemic segmentation of monosyllabic words; the two blending routines jointly support coda-inclusive word construction and continuous synthesis; and Echo Reading extends toward modeled sentence-level prosody. This progression is not prescribed by the source handbook. Although each routine is documented in *Into Reading*, their organization into a single phoneme-to-prosody progression is a conceptual reconstruction by the present authors, drawing on Structured Literacy's principle of systematic, cumulative sequencing rather than an ordering stated in the source. On this basis, the progression supports the Tripartite Adaptation Framework for Scripted Instruction, a Phonemic Segmentation Scaffold, a Coda-to-Word Blending Scaffold, and a Prosodic Oral-Fluency Scaffold, elaborated in the Discussion.

In relation to RQ1, the findings show that all four selected routines demonstrate theoretical compatibility with at least one identified Vietnamese-related barrier, although two require conditional implementation support. In relation to RQ2, the findings establish the documentary basis for translating these routines into bilingual I do – We do – You do teacher-guide scaffolds, which is developed further in the Discussion.

5. Discussion

5.1. From Routine Compatibility to a Micro-Instructional Progression

Read as a staged sequence, the routines suit the Vietnamese context because the three barriers occur at related but distinct levels of oral decoding: phoneme-level segmentation (Tap and Count the Sounds), word-level construction and synthesis (the two blending routines), and connected-speech modeling (Echo Reading). This staging is consistent with Koda's (2005) view that additional-language literacy processing is shaped by established L1 routines and with evidence that English word-final consonants are vulnerable in Vietnamese learner speech (Disney & Le, 2024), linking coda awareness with complete spoken-word formation rather than isolated articulation.

At the connected-speech level, Echo Reading structures imitation of modeled pronunciation and expression, extending support toward phrasing and prosody—supported by evidence that Vietnamese prosodic experience can influence English stress processing (Le, 2021)—though its compatibility depends more on implementation quality. Together, the routines suggest foundational oral fluency may be supported through increasingly complex

demands, and this staged interpretation links the document analysis to the study's material-design contribution.

5.2. Translating Compatibility into Bilingual Scripted Instruction

The findings support the proposed Tripartite Adaptation Framework for Scripted Instruction, reorganizing the four routines around three functions: Phonemic Segmentation, Coda-to-Word Blending, and Prosodic Oral Fluency. The first adapts Tap and Count the Sounds into brief bilingual procedures requiring learners to hear, tap, and identify the phonemes in short words, so practice does not rely on whole-word repetition. The second combines the two blending routines: learners produce the final consonant as part of the word and then retain it while blending continuously—well suited to Vietnamese teacher's guides because it lets teachers anticipate errors such as final-consonant omission or an unintended vowel after a coda. The third adapts Echo Reading for connected production by specifying phrase boundaries, prominent words, and an oral model to imitate, with recorded models or annotated scripts where necessary. Because the framework supports rather than works against Vietnam's ESL-oriented direction, the bilingual prompts are temporary scaffolds: the first-language gloss makes a task instruction unambiguous in a large class, while the target action—segmenting, blending, or imitating—is performed in English. The scripts confine Vietnamese to brief procedural cues and error-monitoring notes rather than the modeled language itself, and these cues can be faded as learners gain familiarity so English becomes the sole language of practice.

Across the three components, the I do – We do – You do sequence provides a practical structure: the teacher models the target action, learners practise with support, then apply it independently while the teacher monitors anticipated L1-related difficulties—consistent with the Gradual Release of Responsibility model (Pearson & Gallagher, 1983; Fisher & Frey, 2014). The implication is not to replace communicative instruction with decoding routines but to strengthen communicative materials by inserting short, targeted routines before or within speaking and oral-reading activities, connecting pronunciation instruction to curriculum design rather than leaving it to unscripted correction (McGregor & Reed, 2018).

The following compact example is an author-constructed illustration of the Coda-to-Word Blending component, not an excerpt from the analysed corpus; because the HMH source is an English-only handbook, the bilingual scripting is part of the proposed adaptation rather than of the documentary data. It illustrates the intended register: short, command-like bilingual prompts suited to large Vietnamese classes, avoiding word-by-word translation. *I do*: "Listen: /s/ … /ĭ/ … /t/ → sit. Watch my mouth on the last sound." (Nghe cô: /s/ … /ĭ/ … /t/. Nhìn miệng cô ở âm cuối.) *We do*: "Blend with me—don't drop the last sound." (Đọc nối cùng cô—đừng bỏ âm cuối.) *You do*: "Your turn. Say the whole word. Is the last sound there?" (Đến lượt em. Đọc cả từ. Có âm cuối chưa?)

5.3. Implications for Primary English Materials and Teacher Support

The analysis has two immediate implications for materials development. First, primary English teacher's guides should make predictable phonological difficulties visible at the level of procedure: instead of directing teachers only to "repeat," "read aloud," or "practise pronunciation," materials should indicate the target sound feature, the expected learner response, and the likely L1-related error to monitor. Second, barriers require different support: phonemic segmentation and word blending can work through simple bilingual prompts, gestures, and selected target words, whereas prosodic modeling is more demanding because its value depends on model accuracy, so materials incorporating Echo Reading should provide

phrase-marked sentences, short audio demonstrations, or teacher preparation notes. If primary learners are to use English increasingly for communication, materials must support the sound-level and prosodic foundations that make oral communication intelligible; the framework adds this support without displacing communicative goals.

Translating these routines into Vietnamese primary classrooms also depends on contextual conditions a document analysis cannot resolve. Teachers' own English pronunciation proficiency is most consequential, since Echo Reading and the prosodic component supply value only insofar as the modeled delivery is accurate, the principal reason recorded models and phrase-marked scripts are proposed. Class size and lesson time constrain implementation: the routines are brief by design, but large classes and crowded schedules mean they must be embeddable in a few minutes rather than treated as separate lessons. Alignment with the existing communicative textbook matters too, because the framework is inserted into current materials rather than replacing them, requiring the routines to attach to words and texts learners already encounter. Finally, age-appropriate engagement is relevant for Grade 1 learners, for whom the tactile and choral character of the routines is an advantage but sustained phoneme-level work must stay short and playful. These conditions do not weaken the theoretical compatibility established above, but they qualify the move to classroom use and define what subsequent piloting would examine.

The contribution must be read within several limitations. First, it analyses four selected routines from one handbook and so does not evaluate the complete HMH programme or Structured Literacy as a whole. Second, it does not compare the routines with currently used Vietnamese primary English textbooks; claims about a need for stronger micro-level decoding support should be understood as a design rationale rather than the result of a completed textbook audit. Third, being qualitative document analysis rather than classroom intervention, it establishes that the routines contain mechanisms theoretically compatible with Vietnamese-related barriers but does not demonstrate that they improve learner production—a distinction especially important for Echo Reading, whose contribution depends on model accuracy, teacher preparation, and implementation conditions.

The next phase should begin with expert review of the bilingual framework by specialists in Vietnamese–English phonology, primary English education, and materials design. Following revision, a small-scale classroom pilot could examine teacher usability, learner participation, and the feasibility of embedding the routines in regular communicative lessons, after which intervention research could evaluate changes in phoneme segmentation accuracy, word-final consonant retention, word-level blending accuracy, and sentence-level oral-reading prosody, alongside teacher usability and student engagement.

6. Conclusion

This study examined whether four foundational Structured Literacy routines from HMH Into Reading Grade 1 are theoretically compatible with three L1-induced phonological barriers facing Vietnamese primary learners—monosyllabic processing, coda deletion, and prosodic transfer—and how compatible routines might be translated into bilingual scripted guidance. The analysis indicates that the four form a coherent phoneme-to-prosody progression, all compatible with at least one barrier and two requiring conditional implementation support. On this basis it proposes the Tripartite Adaptation Framework for Scripted Instruction as an operationally specified, theoretically justified, and explicitly non-validated basis for future testing. Its value lies not in claiming demonstrated learning gains but in offering materials designers a micro-instructional lens for embedding explicit decoding support within Vietnam's ESL-emerging, communicatively oriented classrooms.

References

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Disney, S. J., & Le, N. C. (2024). Word-final /s/–/z/ omission in Vietnamese English. *Languages, 9*(10), Article 327. https://doi.org/10.3390/languages9100327

Fallon, K. A., & Katz, L. A. (2020). Structured literacy intervention for students with dyslexia: Focus on growing morphological skills. *Language, Speech, and Hearing Services in Schools, 51*(2), 336–344. https://doi.org/10.1044/2019_LSHSS-19-00019

Fisher, D., & Frey, N. (2014). *Better learning through structured teaching: A framework for the gradual release of responsibility* (2nd ed.). ASCD.

Houghton Mifflin Harcourt. (2017). *Into reading, Grade 1: Instructional routines*. Houghton Mifflin Harcourt.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge. https://doi.org/10.4324/9780203935927

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524841

Le, G. (2021). L1 Vietnamese L2 English speakers' cues to the perception of stress. *Studies in the Linguistic Sciences: Illinois Working Papers, 2021*. https://www.ideals.illinois.edu/items/118845

McGregor, A., & Reed, M. (2018). Integrating pronunciation into the English language curriculum: A framework for teachers. *The CATESOL Journal, 30*(1), 69–94. https://escholarship.org/uc/item/0cx139qq

Ministry of Education and Training. (2018). *Chương trình giáo dục phổ thông môn Tiếng Anh* [General education curriculum for English] (Promulgated under Circular No. 32/2018/TT-BGDĐT). https://thuvienphapluat.vn/van-ban/Giao-duc/Thong-tu-32-2018-TT-BGDDT-Chuong-trinh-giao-duc-pho-thong-403454.aspx

Nguyen, T. T. T. (2007). *Difficulties for Vietnamese when pronouncing English: Final consonants* [Student thesis, Högskolan i Halmstad]. DiVA Portal. https://www.diva-portal.org/smash/record.jsf?pid=diva2:518290

Nhat, T. N. M., & Hien, H. T. T. (2025). R.E.F.L.E.C.T: Enhancing Vietnamese EFL undergraduates' pronunciation proficiency. *Asian-Pacific Journal of Second and Foreign Language Education, 10*(1), Article 25. https://doi.org/10.1186/s40862-025-00333-0

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524537

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology, 8*(3), 317–344. https://doi.org/10.1016/0361-476X(83)90019-X

Politburo. (2024). *Kết luận số 91-KL/TW về tiếp tục thực hiện Nghị quyết số 29-NQ/TW* [Conclusion No. 91-KL/TW on continuing the implementation of Resolution No. 29-NQ/TW]. https://thuvienphapluat.vn/van-ban/Giao-duc/Ket-luan-91-KL-TW-2024-

tiep-tuc-thuc-hien-Nghi-quyet-29-NQ-TW-doi-moi-can-ban-toan-dien-giao-duc-621156.aspx

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching, 10*(1–4), 209–231. https://doi.org/10.1515/iral.1972.10.1-4.209

#132 EXPLORING AI SELF-EFFICACY IN EFL STUDENTS: IMPLICATIONS FOR LEARNING PERFORMANCE

Phuong Vo Thi Bich[1]

Abstract

The integration of Generative AI (GenAI) has created a paradigm shift in English language learning as it offers transformative benefits, as well as ethical and academic challenges. While previous research focused on how often students use AI tools, a significant gap in understanding how students perceive their capacity to effectively engage with AI tools exists. This study investigated students' AI usage patterns and their confidence in engaging with AI in learning contexts reflected through five key dimensions: AI-assisted learning, anthropomorphic interaction, emotional comfort, technical and prompting skills, and critical evaluation of AI outputs. The study adopted quantitative design. To collect the data, an online questionnaire was administered to 150 English-majored and non-English majored students. The findings indicated that the students demonstrate moderate to high levels of AI self-efficacy, particularly in using AI as a learning support tool. However, lower levels of confidence were observed in more advanced competencies, such as prompt engineering and critical evaluation of AI-generated content. The findings provide empirical insights into students' readiness to engage with GenAI and offer implications for educators to foster more effective, critical, and responsible AI use in language learning.

Keywords: AI self-efficacy, EFL students, academic performance

1. Introduction

In recent years, the rapid integration of generative artificial intelligence (GenAI) has created a paradigm shift in language education. With its various tools such as ChatGPT, Gemini, or DeepSeek, it is believed that the use of GenAI supports learners in enhancing language skills, particularly writing, speaking and communication, as well as reading and

[1] *Phuong Vo Thi Bich, M.A., Faculty of Foreign Languages, HCMC University of Foreign Languages – Information Technology, HCMC, Vietnam; Email: phuongvtb@huflit.edu.vn*

translation. Firstly, studies have shown that AI-powered tools played a central role in assisting learners to improve their writing proficiency through providing instant feedback, helping them generate new ideas, structure writing process, enrich vocabulary (Hartwell & Aull, 2023), and proofread grammar-related mistakes (Waer, 2021; Zhao, 2022). Furthermore, GenAI, especially chatbots, offer a non-judgmental learning environment, and can simulate roleplays which help students enhance their speaking fluency and pronunciation (Yuan & Lyu, 2024; Ismail & Alharkan, 2024; Kittredge et al., 2025; Dong et al., 2022). In addition, GenAI enables simplifying texts and translating complex information between different languages effortlessly, leading to improvements in reading and translation skills (Celik et al., 2025; Wang & Fussell, 2026).

In addition, because AI-generated feedback is timely and well-structured, it would give students a sense of task completion which significantly enhances students' self-efficacy. In this context, self-efficacy serves as a mediating factor and contributes to the behavioral intentionality of students and leads to the academic success (Zhou & Wang, 2026; Shahat et al., 2025; Qin et al., 2026). In addition, the use of AI applications can create a supportive and low-stress learning environment, which may help lower foreign language anxiety (Yuan & Lyu, 2024; Pertiwi et al., 2025; AlTwijri & Alghizzi, 2024). On the other hand, through the implementation of GenAI tools, personalization of learning and motivation for studying will be achieved (Liu et al., 2026).

However, although there were so many benefits considering the utilization of GenAI, many scholars also identified some potential drawbacks. Firstly, excessive dependency on AI-driven technologies to complete tasks like writing and translating with ChatGPT will make learners metacognitively lazy and, hence, hinder their critical and creative thinking, problem solving, and independent reasoning (Sardi et al., 2025; Zhai et al., 2024). Secondly, the utilization of AI-driven technologies was associated with several ethical concerns like cheating, plagiarism, bias in algorithms, and difficulty in capturing cultural nuance (Almusharraf et al., 2025; Wang & Fussell, 2026; Bećirović et al., 2025).

In the context where AI tools are increasingly accessible and their benefits and risks become more evident, AI self-efficacy has become a crucial factor in determining whether students can engage with these tools meaningfully and achieve academic success. AI self-efficacy refers to students' internal belief in their ability to use them effectively (Rajapakse et al., 2024). In essence, AI self-efficacy affects the decision to perform an activity, the degree of commitment to the activity, and the level of persistence when faced with difficulties (Xia, 2026). Highly efficacious learners treat difficult tasks as challenges and not obstacles. They also actively employ AI tools, set themselves high goals, and monitor their learning process (Gao et al., 2025; Bećirović et al., 2025).

In addition, regarding tasks such as writing, students with high AI self-efficacy demonstrate greater readiness in searching for, analyzing, and internalizing feedback from artificial intelligence systems (Thurman, 2026). Rather than being passive in the face of AI-generated outputs, they use it as a tool to improve themselves and exhibit stronger feedback literacy and autonomy (Almusharraf et al., 2025). Moreover, AI self-efficacy has been shown to have a direct impact on one's educational achievement. There exists a significant and direct correlation between an individual's academic self-efficacy and his or her learning outcomes in AI-mediated settings (Xu et al., 2024). In combination with an appropriate environment, it improves learners' proficiency and task performance (Zhou & Wang, 2026; Shahat et al., 2025).

While the adoption of AI technology may create stress for learners, self-efficacy serves as a strong coping resource. Students with high self-efficacy tend to experience considerably lower levels of learning anxiety when using AI technology (Tang et al., 2026; Qin et al., 2026). By alleviating fear and emotional distress, self-efficacy contributes to freeing up cognitive capacities, enabling students to engage in deep learning and avoid stress (Xu et al., 2024; Yan et al., 2025). Lastly, one of the risks related to the use of AI is its negative effects on learning independence. High levels of dependency on AI might cause students to depend excessively on the technology to complete academic tasks instead of doing so themselves. Nevertheless, those students with high AI self-efficacy adopt a proactive attitude towards the use of AI and utilize it to achieve their learning goals (Zhang et al., 2024). At the same time, those students whose self-efficacy is low and who doubt their own capabilities resort to AI as a compensatory strategy, which results in the risk of dependency (Huang & Mizumoto, 2024).

Considering the potentials and challenges of AI tools integration for learners and the gaps in previous studies, which mainly investigated AI acceptance, AI usage patterns, or skills-related outcomes, this study focuses on students' perceived capacity to use AI tools across different levels from basic learning support to prompting and critical evaluation. This study considers AI self-efficacy as a multidimensional construct and aims to examine which dimensions have the strongest impact on students' academic performance.

This study was conducted for the purpose of exploring the relationship between AI self-efficacy and academic performance in EFL contexts. It aimed at examining the current levels of AI self-efficacy among EFL students across five specific dimensions: AI-assisted learning, anthropomorphic interaction, emotional comfort, technical and prompting skills, and critical evaluation. Secondly, it seeks to evaluate the predictive power of AI self-efficacy on the academic performance of EFL students.

The current study seeks to answer the following research questions:

To what extent do EFL students demonstrate AI self-efficacy across the five dimensions?

To what extent does AI self-efficacy significantly predict the academic performance of EFL students?

2. Literature review

2.1. Defining AI self-efficacy in education

The term AI self-efficacy has been described as "a person's sense of efficacy to accomplish certain activities or go through certain scenarios involving AI technologies" (Zhang et al., 2025, p. 4). It refers to the confidence and assurance of a learner concerning their ability to effectively arrange, undertake, and engage with the use of AI technologies to accomplish a set goal (Gao et al., 2025; Yuan & Lyu, 2024). Rooted in Bandura's social cognitive theory, which proposes that self-efficacy determines the actions of a person including motivation, effort, and persistence, this concept has been modified and tailored to meet the requirements of current AI (Zhang et al., 2025). In education settings, AI self-efficacy refers to students' belief in themselves to independently work with AI technologies and learn how to solve their academic problems efficiently (Zhang et al., 2025; Gao et al., 2025; Cong-Lem et al., 2025). For instance, in writing, AI self-efficacy entails the degree to which students believe in their capability to use AI technologies to write successfully (Cong-Lem et al., 2025; Shen & Tao, 2025). A student with high levels of AI self-efficacy is more likely to be proactive while

working with AI technologies. They are able to cope with difficulties in their studies and experience less classroom and learning-related anxiety (Yuan & Lyu, 2024; Gao et al., 2025).

2.2. Dimensions of AI self-efficacy

Previous studies have characterized AI self-efficacy as a multidimensional construct that covers operational, cognitive, affective and metacognitive dimensions. The first dimension of AI self-efficacy reflects learners' confidence in interacting with AI tools. It involves learners' technical and prompting skills as well as their understanding of how AI functions, its mechanisms, and the cognitive processes behind AI tools (Bećirović et al., 2025). Cognitive dimension reflects learners' confidence in utilizing AI as an intellectual partner to achieve learning goals. This dimension includes the use of AI to support learning, research, and self-regulation activities such as goal setting, planning and monitoring progress (Shen & Tao, 2025). The affective dimension of AI self-efficacy comprises comfort and anthropomorphic interaction, which reflects learners' beliefs in their own capability to effectively interact with AI agents that imitate human dialogue. This dimension requires in learners necessary resilience to use AI without technological stress (Wang & Chuang, 2024). Last but not least, metacognitive and ethical dimension ensures that learners use AI in a productive and ethical manner. It requires critical evaluation of AI outputs and learners' ability to detect factual inaccuracies, algorithmic biases, and hallucinations (Almusharraf et al., 2025). In addition, ethical and responsible integration of AI tools into learning is also required because it reflects learners' confidence to navigate ethical dilemmas such as plagiarism, data privacy and ensures that learners use technology within established boundaries (Bećirović et al., 2025).

Other researchers also suggest context-specific dimensions such as productive and receptive efficacy, or writing efficacy (Liu et al., 2026; Shen & Tao, 2025). For instance, for productive and receptive efficacy, learners distinguish between their confidence in using AI tools for active output of speaking and writing skills or comprehension tasks such as listening and writing (Liu et al., 2026). On the other hand, for academic writing, self-efficacy is broken down into smaller steps in which AI tools are involved in the learning process such as brainstorming, linguistic rules, and self-regulation (Gao et al., 2025).

Within the current study, AI self-efficacy is referred to as a multidimensional construct encompassing five main dimensions of AI assistance, anthropomorphic interactions, comfort, technical and prompting skills, and evaluation of AI-generated outputs. These dimensions reflect not only whether a learner uses AI, but also how effectively, confidently, and ethically they integrate it into their own learning pathway.

2.3. The role of AI self-efficacy in language education

The literature demonstrates consistent positive relationships between the level of AI self-efficacy and students' academic results and output quality in English courses. Specifically, academic self-efficacy among EFL learners working in an AI-assisted environment positively impacts their English learning achievements to a high degree with the maximum value in one of the structural equation models is equal to 0.940 (Xia, 2026). Students who believe in their abilities to use various technological tools including the generative AI software are employing particular writing strategies based on their use. These actions improve L2 writing proficiency and satisfaction levels among EFL students.

Moreover, AI self-efficacy acts as an engine of learner motivation and multidimensional engagement through behavioral, cognitive, and affective processes. According to self-determination theory, using AI-based tools that are customized for one's needs allows for the fulfillment of basic psychological needs for competence, thus stimulating

autonomous, intrinsic motivation in the learner (Liu et al., 2026). For instance, using ChatGPT in an organized manner within EFL classrooms has been found to positively affect students' ideal L2 self and L2 experience motivations, resulting in enhancing the relationship between writing self-efficacy and motivational orientations of learners (Huang & Mizumoto, 2024). Moreover, self-efficacy is related to learning grit; those who show more grit will have more mastery experiences in using AI tools, thus enhancing their self-efficacy.

Finally, within the framework of the technology acceptance model (TAM), self-efficacy is an essential construct that links perception of the AI-based tools by the learners to their actual behavior and foreign language learning. Specifically, the higher EFL students perceive the generation AI-based tool to be as easy and useful to employ, the stronger will be their self-efficacy in terms of the tool use, which, in turn, will predict their intention to use it further (Almusharraf et al., 2025).

2.3. The relationship between AI self-efficacy and academic performance

Previous studies have revealed an obvious correlation between AI self-efficacy and academic performance in which academic performance can be predicted by AI self-efficacy (Teng, 2025). As seen from the findings of the study by Xia (2026) who used structural equation modeling in researching English as a Foreign Language (EFL) students, there was a huge path coefficient (β=0.940) of students' academic self-efficacy with regard to their learning outcome when utilizing generative AI technology. High AI self-efficacy enables students to be certain about their capability to reach academic goals, hence putting them in a position to make necessary efforts and persevere in the learning process using AI technologies (Zhang et al., 2025). It has been similarly noted that students with high self-efficacy predict positive test outcomes (Hao et al., 2025).

However, important cautions when considering how one's AI self-efficacy impacts their academic performance in general are also noticed. Whereas Bećirović et al. (2025) demonstrated that one's AI self-efficacy indeed plays a significant role in predicting the level of quality of the generated AI output (β=0.229), it appears that its influence on the overall academic performance of students is not as significant and can be considered merely positive, yet statistically insignificant (Bećirović et al., 2025). An important risk should be noted in this context. If the application of students' AI self-efficacy results only in producing prepackaged outputs, and does not involve necessary mental operations like analyzing, reading, questioning, and problem-solving, then AI usage could have a detrimental effect on learning outcomes (Xia, 2026). This is because although students might seem highly self-confident in the use of AI technology, their self-confidence is only aimed at acquiring outputs.

2.4. Conceptual framework

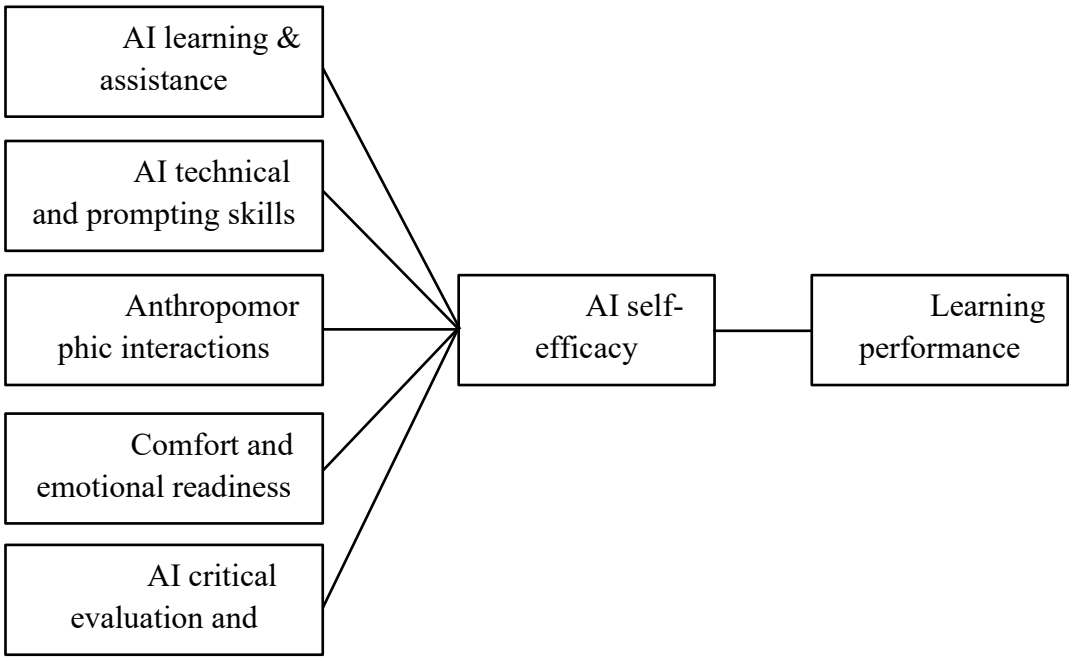
The conceptual framework of AI self-efficacy in the current study encompasses five components: AI learning and research assistance, AI technical and prompting skills, anthropomorphic interactions, comfort and emotional readiness, and technical and prompting skills, AI critical evaluation and responsible use. Firstly, AI technical and prompting skills represent the foundational, technical confidence required to navigate AI systems effectively. It includes not only learners' technological understanding but also their practical application of AI tools. Secondly, anthropomorphic interactions measure learners' comfort and effectiveness in treating AI as a conversational partner, virtual tutor, or collaborative peer. Thirdly, AI learning and research assistance reflects learners' confidence in using AI as an intellectual partner to achieve specific academic and research tasks. Fourthly, the affective dimension of AI self-efficacy is characterized by learners' comfort and emotional readiness associated with

AI usage. Highly efficacious learners perceive AI as a non-judgmental partner, which creates a psychologically safe environment that significantly reduces foreign language anxiety and fear of negative evaluation and increases learners' perceived control of AI tools as well as enjoyment. Finally, AI critical evaluation and responsible use reflect learners' confidence in applying higher-order thinking to review, question, and interpret AI-generated outputs. Efficacious users trust their ability in detecting algorithmic biases, factual inaccuracies. In addition, this dimension also measures learners' ability to resist dependency on AI, in which learners tend to blindly accept AI feedback or use it as a shortcut to replace genuine learning. Another critical characteristic of AI efficacious learners is reflected in their confidence in navigating ethical decisions related to plagiarism and data privacy and the use of technology in a responsible way within academic environments.

This conceptual framework was adapted from Wang and Chuang's (2024) four-dimensional framework comprising assistance, anthropomorphic interaction, comfort, and technological skills. However, an important fifth dimension, AI critical evaluation and responsible use, was added to ensure learners' confidence leads to efficient and ethical integration of AI tools into learning. In this conceptual framework, AI technical and prompting skills serve as the foundational engine. Once a learner feels technically capable, their comfort and emotional readiness are activated, which in turn facilitates smooth anthropomorphic interaction, lowering anxiety and making the technology feel accessible and safe. These foundational and affective dimensions empower learners to utilize AI for deep learning and research assistance. But in order for this aid to translate into real academic success, the whole process must be guided by the critical analysis and responsible use of AI. In total, these five components make up an AI self-efficacy framework, determining not only the frequency of AI tool usage, but also the quality of its correlation with successful academic outcomes.

Figure 2.1.

A conceptual framework of AI self-efficacy (adapted from Wang and Chuang, 2024)



3. Methodology

3.1. Pedagogical settings & participants

This study was conducted with the participation of 150 students from HCMC University of Foreign Languages – Information Technology. The students were currently attending different courses in the second semester of the academic year 2025-2026. Convenience sampling was used for the current study. Below is a table describing the participants' demographic information.

Table 3.1.

The participants' demographic information (n= 150)

Variables		**Number**	**Percentage**
Gender	Male	36	24%
	Female	114	76%
Major	English	114	76%
	Non-English	36	24%
Year	First year	112	74.7%
	Second year	38	25.3%
Scores	Excellent	72	48%
	Good	37	24.7%
	Fair	31	20.7%
	Poor	10	6.7%

Demographic characteristics of participants were provided in Table 3.1. There were totally 150 participants in this sample. With respect to gender, female participants comprised 76%, while male participants constituted 24% of the sample. English-majored students comprised 76%, while non-English-majored participants comprised 24%. In regard to academic year, freshmen comprised 74.7%, while sophomores comprised 25.3% of the participants. In terms of academic performance, most participants performed excellently. They constituted 48% of the sample. Those who performed very well and well comprised 24.7% and 20.7%, respectively. Poor performers comprised only a minor portion of the participants (6.7%). Overall, the sample largely comprised female, English-majored freshmen with high academic performance. Although the sample is heavily weighted toward female participants and English language majors, this imbalance indeed reflects the genuine demographic characteristics of university level EFL programs.

3.2. Design

This study employed a quantitative research design utilizing a cross-sectional survey methodology. This approach was chosen over qualitative or mixed-methods design because it's suitable for quantifying the perceptions of behaviors of a large student population, which ensure objectivity. In addition, by using a structured survey, this study can statistically analyze how specific dimensions of AI self-efficacy interact within a large number of participants. Specifically, quantitative data was collected through an online survey with the participants. The purpose of the survey is to measure the current level of AI self-efficacy of the participants. Quantitative data was collected via an online questionnaire delivered through the LSM Moodle system of the university. The questionnaire was adapted from the framework designed by Wang and Chuang (2024), with the addition of a fifth dimension incorporated based on

literature review of the current literature about AI literacy. The questionnaire included 44 items and was divided into three main parts. The first part aimed at collecting students' demographic information and their academic performance characterized by the midterm scores of the current subjects. The second part investigated the students' current level of AI usage including frequency, purposes, and perceived effectiveness and AI usage pattern including using AI as writing assistance, learning support and analytical tasks. The third part measured students' AI self-efficacy focusing on the five dimensions of AI including learning and research assistance, AI technical and prompting skills, anthropomorphic interactions, comfort and emotional readiness, and AI critical evaluation and responsible use.

Reliability scores of the questionnaire are presented on Table 3.2 below:

Table 3.2.

Reliability of the instrument

Dimension	Cronbach's Alpha	Number of items
Learning and research assistance	0.834	5
AI technical and prompting skills	0.626	2
Anthropomorphic interactions	0.853	5
Comfort and emotional readiness	0.870	6
AI critical evaluation and responsible use	0.836	4

In addition, midterm grades used in the current study were self-reported by the students, based on the scores provided by their instructor. Given that these grades originate from different courses, threat to internal validity is mitigated by a highly standardized grading structure across all courses. Specifically, midterm scores comprise a uniform weight distribution consisting of 10% attendance, 30% self-study, and 60% progress tests.

3.3. Data collection & analysis

Quantitative data were analyzed utilizing statistical tests on SPSS 26. Descriptive statistics in terms of frequency, percentage, mean, and standard deviation were used for summarizing demographic information about participants, extent of use of AI technology, and levels of AI self-efficacy. Correlation analysis via Pearson correlation was used to analyze the relationship between self-efficacy and academic performance. In addition, linear regression was used for analyzing the predictive effect of self-efficacy with regard to academic performance of the learners.

4. Findings

4.1. AI usage pattern among the participants

Table 4.1.1.

Students' purposes of AI use

Purposes of use	Frequency	Percentage
Writing / correcting grammar / editing English texts	129	86%
Summarizing learning materials	110	73.3%
Practicing speaking and pronunciation	51	34%

Translating	85	56.7%
Finding ideas or outlines for essays / presentations	123	82%
Doing quizzes / exam preparation	33	22%

Table 4.1.1 shows the goals students have in mind when using AI technologies in their English studies. The first goal mentioned by 86% of the respondents is associated with the usage of AI in the processes of writing, proofreading, and editing their English texts. Thus, it can be concluded that students use AI mostly as a means of developing their writing skills. Another goal of using AI tools by students is connected with finding some ideas and creating outlines for essays or presentations. The percentage of students who use AI in this case was equal to 82%, which demonstrates that artificial intelligence technologies can play a vital role during the stage of preparing an essay or presentation. It should be added that 73.3% of students used AI to create summaries of educational materials. Furthermore, over 50% of the participants utilized AI for translating English text into Russian, which can also help students learn the language better. However, there were fewer students who used artificial intelligence for practicing speaking skills and pronunciation (34%) and preparing for exams/quizzes (22%). Overall, the results indicate that students employ AI primarily for text-related and academic purposes, which include improving writing and studying course materials. Consequently, AI can be considered a tool for generating and understanding languages rather than a tool for practicing speaking skills.

Table 4.1.2

Perceived effectiveness, dependency and revision degree (n = 150)

Dimensions	**Levels**	**Frequency**	**Percentage**
Frequency of use	Sometimes	2	1.3%
	Often	74	49.3%
	Very often	74	49.3%
Perceived effectiveness	Moderately effective	28	18.7%
	Effective	68	45.3%
	Very effective	54	36%
Perceived dependency	Not dependent at all	2	1.3%
	Slightly dependent	10	6.7%
	Moderately dependent	71	47.3%
	Dependent	63	42%
	Highly dependent	4	2.7%
Revision degree	Minor revision needed	8	5.3%
	Moderate revision needed	63	42%
	Substantial revision needed	62	41.3%
	A lot of revision needed	17	11.3%

Table 4.1.2 reflects the opinions of the respondents about their AI usage for learning English. Generally speaking, students reported quite a high frequency of using AI. All but one

respondent used AI tools often or very often, whereas most participants used AI very often. As far as perceived effectiveness is concerned, the majority of respondents believed that AI tools were effective or very effective for learning English. It means that the majority of respondents had a good impression about the utility of using AI in the process of learning. When it comes to the degree of dependence, the majority of students reported that they had a moderate or relatively high degree of dependence on AI. The biggest percentage was reported for moderately dependent students. Next, there was a high percentage of dependent students. Considering the degree of revision is concerned, it is possible to note that most students reported a need for moderate or considerable revision of the results produced by AI.

4.2. AI self-efficacy among the participants

Table 4.2.

Students' AI self-efficacy (n=150)

Dimensions	**Mean**
AI learning & research assistance	4.20
Anthropomorphic interaction	3.85
Comfort & emotional readiness	4.11
AI technical and prompting skills	3.05
AI critical evaluation and responsible use	2.90
AI self-efficacy	3.62

Table 4.2 reveals students' level of AI self-efficacy. As can be seen, the students' overall AI self-efficacy level is at the mean score of 3.62. Based on the interval classification scale (1.00 – 1.80 Very Low; 1.81 – 2.60 Low; 2.61 – 3.40 Moderate; 3.41 – 4.20 High; 4.21 – 5.00 Very High), this level of confidence is categorized at the high level. Among the five dimensions of AI self-efficacy, AI learning & research assistance occupies the highest rank with the mean score of M = 4.2 (high level). This result shows that students are confident in using AI as a support tool to solve learning problems and increase their productivity.

Comfort & emotional readiness and anthropomorphic interaction rank second and third with M = 4.11 (high level) and M = 3.85 (high level). Students are confident in communicating with AI as if it were a human through conversational flow. In addition, they also demonstrated a high level of emotional ease and reduced anxiety when interacting with AI tools. In addition, students' AI technical and prompting skills are at an average level with M = 3.05 whereas AI critical evaluation and responsible use rank last with the mean score of 2.90 (average level). This lower level reveals that students often struggle with complex prompts and have difficulties judging the accuracy, bias, and reliability of AI-generated contents.

In general, while students demonstrate high levels of confidence in utilizing AI for basic learning and research assistance, their self-efficacy regarding critical evaluation and responsible use is noticeably lower. This suggests a significant gap in terms of confidence among the students. This is because while the students were very confident, comfortable, and prepared to use AI tools effectively for performing learning tasks, their self-assurance about their ability to use AI technology and prompting engineers were not at the same level.

4.3. The relationship between AI self-efficacy and academic performance

Table 4.3.1.

Correlation between AI self-efficacy and academic performance

Correlation		AI learning & research assistance	Anthropomorphic interaction	Comfort & emotional readiness	AI technical and prompting skills	AI critical evaluation and responsible use
Scores	Pearson correlation	-.152	.13	-.040	.285**	.340**
	Sig. (2-tailed)	.064	.878	.624	.000	.000
	N	150	150	150	150	150

The correlation between AI self-efficacy dimensions and student's academic performance is represented in Table 4.3.1. It can be observed that there were only two significant correlations found between the AI self-efficacy dimensions and student's scores – the dimension AI technical and prompting skills and the dimension AI critical evaluation and responsible use. Thus, AI technical and prompting skills demonstrated a positive correlation, which indicates that those students who felt more self-assured in writing prompts and utilizing AI in technical ways obtained higher academic scores. AI critical evaluation and responsible use was positively correlated with higher student's performance, which means that students who were able to critically evaluate the material generated by AI performed better.

However, AI learning and research support, anthropomorphic engagement, and emotional preparedness did not have any statistical association with the performance of the students. It means that only finding AI useful, comfortable interactions with AI, and being emotionally prepared to use AI alone cannot contribute to higher achievement at school. All in all, the results seem to indicate that there is a stronger link between better academic achievements and students' high-level AI competencies, namely, prompting and critical assessment.

Table 4.3.2.

Results of multiple regression analysis examining the effects of AI self-efficacy dimensions on academic performance

Variables	B	β	t	Sig
AI learning & research assistance	-0.278	-0.136	-1.879	.062
Anthropomorphic interaction	-0.010	-0.006	-0.077	.939
Comfort & emotional readiness	-0.128	-0.063	-0.867	.387
AI technical and prompting skills	0.400	0.321	4.395	.000***
AI critical evaluation and responsible use	0.456	0.389	5.303	.000***

Statistic	Value
R	0.497
R^2	0.247
Adjusted R^2	0.221

F	9.451
Sig. (F)	.000***
N	150

*Note: *p < .05, **p < .01, ***p < .001*

The effects of self-efficacy on learning outcomes have been investigated using linear regression analysis in Table 4.3.2 below. The linear regression model has been proven to be statistically significant at F=9.451, p<.001 with $R^2 = 0.247$ or an explanatory power of 24.7% in predicting students' performance. The regression analysis indicates that general AI self-efficacy does not predict academic success. Instead, its impact is dependent on specific, high-level competencies. Out of the five factors considered, technical prompting skills and critical evaluation turned out to be important positive predictors of academic performance, while help, anthropomorphism, and comfort failed to be statistically significant. More precisely, technical proficiency and ability to engineer prompts played a positive part in the influence exerted on the dependent variable ($\beta = 0.321$, $p < .001$), which means that the greater technical expertise and proficiency is in terms of creating appropriate prompts, the higher dependent variable becomes. It is worth noting that the AI critical evaluation and use of AI turned out to be the strongest predictor of this outcome ($\beta = 0.389$, $p < .001$), which demonstrates the importance of critical evaluation and use in achieving high academic performance. In contrast, such variables as help, anthropomorphism, and emotional comfort did not prove to be statistically significant at all, giving p-values of .062, .939, and .387 respectively.

5. Discussion

The present study was carried out with the aim of studying students' AI self-efficacy level and its relationship with learning performances. According to the findings, there are several important points which could be highlighted.

First of all, in relation to students' interaction with AI for learning purposes, the findings reveal that students engage with AI tools at a high level across writing, learning support, and analytical tasks, confirming that AI is no longer a novelty but a routine academic instrument. These findings are consistent with the results obtained by Cong-Lem et al. (2025), who investigated English-majored students in Vietnam, also reported a high level of interaction with AI (42% use it regularly, 8% use AI daily whereas 45% use it occasionally). Besides, Fauzi et al. (2025) states that students' interactions with AI are positively correlated with their attitudes and willingness to use it in the future. In general, the results point to a pattern in which students use AI more proficiently for improving writing skills and facilitating learning processes rather than using it as an aid for critical reasoning. Moreover, these findings correspond to the trend within higher education where students are actively using AI not only for personal assignments such as helping with writing but also entering into an ecosystem of various tools that they use frequently to achieve learning goals.

Secondly, the results show that the students had a relatively high level of self-efficacy with respect to AI, meaning that they believed that they were able to use AI tools in order to support their learning and academic activities. Similar results were found in the study conducted by Almusharraf et al. (2025), who claimed that EFL students in Saudi Arabian universities demonstrated high levels of AI self-efficacy with M= 3.77. Particularly, the students expressed high levels of self-assurance in using AI for idea generation and improving writing skills. In addition, the same pattern was observed in the research conducted by Yuan and Lyu (2024), which involved Chinese students. However, there is still evidence of an existing confidence gap. Students were generally confident in the use of AI to facilitate general

learning but were only moderately confident about their technical and prompting competencies, with the lowest being related to critical evaluation and responsible usage. It is evident that students have confidence in their basic knowledge and application of AI technology rather than more complex competencies, including the ability to develop prompts, control output, spot bias, and evaluate AI-generated content for accuracy. The findings mirror those in a similar study conducted by Hao et al. (2025). Despite their positive perception of AI as beneficial and easy to utilize, the self-efficacy levels of Chinese university students stayed relatively low. In addition, the use of GenAI more often was not linked to higher levels of confidence. As stated in the study, the reason for this phenomenon may lie in students' poor exposure to AI technologies, access limitations, and lack of institutional guidance on proper integration of AI technology in learning processes.

Lastly, the findings from the regressions indicate that academic performance is associated with AI self-efficacy, specifically with higher-order competence of AI technical and prompting skills and AI critical evaluation and responsible use within AI self-efficacy. This result aligns with Xia's (2026), who studied the effects of self-efficacy on the academic achievements of students in relation to generative AI. Furthermore, the same result was obtained by Hao et al. (2025) during their research on the connection between self-efficacy and predictions of students' academic performance in tests. Therefore, the results of the current study can be confirmed by the above literature evidence. This means that the students who have higher self-efficacy are more inclined to be successful in their studies due to the high motivation for mastering AI tools.

On the other hand, it becomes clear that not all dimensions of AI self-efficacy affect the same way. Indeed, academic outcomes depend on students' technical competence and prompting abilities, alongside critical evaluation and proper utilization of AI-generated text. Thus, merely perceiving AI technology comfortably and as an assisting tool cannot lead to improvement of learning outcomes. The key point here refers to the possibility of controlling AI interactions, generating appropriate prompts, evaluating AI-generated outputs, and making ethical decisions while working with AI. In such a manner, it seems that academic success may differ from the frequency of AI utilization. In fact, it was demonstrated that high self-efficacy positively influences interaction with AI technologies, although frequent utilization does not necessarily result in successful learning.

Overall, it is important for educators to realize the major pedagogical implications of the results that students must not just be encouraged to use AI in their studies, but also taught how to use AI wisely, effectively, and ethically. If learners are able to take their AI-based activities to a higher level through developing competencies such as prompting and evaluating, their learning experience becomes far more valuable.

It follows from the results of this study that AI self-efficacy plays a significant role in motivating students to work with technologies based on artificial intelligence while high-level competencies such as technical prompting and critical evaluation play an even more important role in influencing academic performance. These results show that it is crucial for educators to encourage their learners to engage in reflective activities related to AI use in the educational process. In other words, learners' AI self-efficacy can contribute to improved academic achievement only when confidence is accompanied by meaningful engagement with AI. Students should use AI as a collaborative learning tool to support reasoning, reflection, revision, and knowledge construction rather than merely to produce ready-made answers.

6. Conclusion & recommendations

In conclusion, the present study shows that the use of AI tools by EFL students is characterized by sufficient confidence. Furthermore, academic success is mostly associated with the ability to utilize technology and prompt writing, combined with critical assessment and proper application. This implies that it is the ability of students to create appropriate prompts, assess, correct, and validate AI output, as well as apply AI in an ethical manner that is likely to make them succeed academically. On the other hand, being familiar with AI and considering AI to be a helpful assistant is unlikely to lead to higher academic success.

From the results, some important implications were drawn. Firstly, the major implication for students is that it is not enough to become proficient in using AI often; they should focus on how to use AI effectively. Specifically, this involves critical analysis and verification of what has been produced by AI rather than copying AI suggestions without assessing and improving them. With the finding that education must move beyond simply giving students access to AI technologies, instructors must directly educate students on appropriate prompting techniques, AI output evaluation, and ethical use of AI. Students should be required in their assignments to detail their use of AI, modifications made, and evaluations of the output quality.

For educational institutions, the implications are that they must move beyond just providing AI technology access to developing students' AI competency. This involves creating AI competency frameworks, embedding AI literacy into educational programs, and establishing guidelines for responsible use of AI. Institutions should also focus on educating educators about AI pedagogy and assessment using AI so that it can be utilized responsibly.

In summary, mere access to AI technology is not sufficient in terms of its effective integration in EFL learning. Students require the necessary competency to be able to employ AI technology in an effective manner.

References

Almusharraf, A., Bailey, D., Almusharraf, N., & Alotaibi, T. (2025). Students' perceptions of generative AI in EFL writing: Strategies, self-efficacy, satisfaction and behavioural intention. *Australasian Journal of Educational Technology*, *41*(5), 18-36. https://doi.org/10.14742/ajet.10045

AlTwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, *10*(10). https://doi.org/10.1016/j.heliyon.2024.e31053

Bećirović, S., Polz, E. & Tinkel, I. (2025). Exploring students' AI literacy and its effects on their AI output quality, self-efficacy, and academic performance. *Smart Learning Environment,* 12, Article 29. https://doi.org/10.1186/s40561-025-00384-3

Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, *8*(1), 102418-102418. https://doi.org/10.29140/ajal.v8n1.102418

Dong, L., Jamal Mohammed, S., Ahmed Abdel-Al Ibrahim, K., & Rezai, A. (2022). Fostering EFL learners' motivation, anxiety, and self-efficacy through computer-assisted

language learning-and mobile-assisted language learning-based instructions. *Frontiers in psychology*, 13, 899557. https://doi.org/10.3389/fpsyg.2022.899557

Fauzi, C., Rahmani, E. F., & Utimadini, N. J. (2025). Beyond technology acceptance: An interplay of self-efficacy, language proficiency, and ChatGPT adoption from a TAM perspective. *Jurnal Pendidikan Progresif*, *15*(3), 1970-1988. https://doi.org/10.23960/jpp.v15i3.pp1970-1988

Teng, M. F. (2025). Metacognition, Self-efficacy Belief, Language Learning Motivation, and Perceived Progress in Online English Learning: A Cross-lagged Analysis. *International Journal of TESOL Studies*, *7*(1). https://doi.org/10.58304/ijts.20250102

Gao, J., Zhang, J., & Li, Y. (2025). Do AI Chatbot-integrated writing tasks influence writing self-efficacy and critical thinking ability? An exploratory study. *Computers and Education: Artificial Intelligence*, 100472. https://doi.org/10.1016/j.caeai.2025.100472

Hao, T., Zhao, X., & Lu, Z. (2025). Spontaneous use of generative artificial intelligence in preparing for college English test: GenAI adoption factors, learners' motivation, and self-efficacy. *Technology in Language Teaching & Learning*, *7*(3), 102640-102640. https://doi.org/10.29140/tltl.v7n3.102640

Hartwell, K., & Aull, L. (2023). Editorial Introduction - AI, corpora, and future directions for writing assessment. *Assessing Writing, 57,* Article 100769. https://doi.org/10.1016/j.asw.2023.100769

Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics*, *1*, 102324-102324. https://doi.org/10.29140/dal.v1.102324

Ismail, S., & Alharkan, A. (2024). EFL learners' positive emotions in the era of technology: Unpacking the effects of artificial intelligence on learning enjoyment, self-efficacy, and resilience. *Computer-Assisted Language Learning Electronic Journal*, *25*(4), 526-551. https://callej.org/index.php/journal/article/view/471

Kittredge, A. K., Hopman, E. W., Reuveni, B., Dionne, D., Freeman, C., & Jiang, X. (2025). Mobile language app learners' self-efficacy increases after using generative AI. In *Frontiers in Education* (Vol. 10, p. 1499497). Frontiers. https://doi.org/10.3389/feduc.2025.1499497

Cong-Lem, N., Bui, C. N. V., Nguyen, N. P. N., & Nguyen, P. T. M. (2025). Exploring vietnamese EFL learners' generative AI usage patterns, self-efficacy and self-directed learning: preliminary findings. The 13th OpenTESOL International Conference 2025. https://doi.org/10.26180/29192639.v3

Liu, J., Liu, M., Yao, Y., & Li, D. (2026). Perceptions and Acceptance of Generative Artificial Intelligence Influencing Chinese EFL Learners' Engagement in Informal Digital Learning of English: Mediating Roles of Self-Efficacy and Motivation. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.70122

Liu, Z. M., Hwang, G. J., Chen, C. Q., Chen, X. D., & Ye, X. D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 466-490. https://doi.org/10.1080/09588221.2024.2389923

Pertiwi, P. C., Segoh, D., & Rohmadhani, A. (2025). Artificial Intelligence in Academic Environments: Reducing or Increasing Foreign Language Anxiety?. *Inspiring: English Education Journal*, *8*(1), 100-131. https://doi.org/10.35905/inspiring.v8i1.13009

Qin, F., Liu, B., Li, W., Li, H., & Li, J. (2026). The Mediating Role of Motivation Between Generative Artificial Intelligence and Foreign Language Proficiency: A Technology Acceptance Model Perspective. *Sage Open*, *16*(1), 21582440261430425. https://doi.org/10.1177/21582440261430425

Rajapakse, C., Ariyarathna, W., & Selvakan, S. (2024). A self-efficacy theory-based study on the teachers' readiness to teach artificial intelligence in public schools in Sri Lanka. *ACM Transactions on Computing Education*, 24(4), 1–25. https://doi.org/10.1145/3691354

Sardi, J., Darmansyah, Candra, O., Yuliana, D. F., Habibullah, Yanto, D. T. P., & Eliza, F. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? a systematic review. *International Journal of Engineering Pedagogy* (*iJEP), 15*(1), 94–108. https://doi.org/10.3991/ijep.v15i1.53379

Sari, E., & Han, T. (2024). The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. *Journal of Computer Assisted Learning*, *40*(5), 2065-2080. https://doi.org/10.1111/jcal.13004

Shahat, H. A., Badawy, M. E., Elballah, K. A., & Ibrahim-Shook, A. (2025). Self-efficacy as a mediator between ChatGPT usage and research motivation among postgraduate students. *International Journal of Innovative Research and Scientific Studies*, *8*(4), 987-996. https://doi.org/10.53894/ijirss.v8i4.7982

Shen, X., & Tao, Y. (2025). Metacognitive Strategies, AI-based Writing Self-efficacy and Writing Anxiety in AI-assisted Writing Contexts: A Structural Equation Modeling Analysis. *International Journal of TESOL Studies*, *7*(1), 70–87. https://doi.org/10.58304/ijts.20250105

Tang, J., Zhang, A., Sun, M., Leng, X., & Luo, L. Drivers of Willingness to Communicate with Generative AI: The Roles of Self-Efficacy, Grit, Speaking Enjoyment, and Anxiety from a Self-Determination Theory Perspective. *Frontiers in Psychology*, *17*, 1754495. https://doi.org/10.3389/fpsyg.2026.1754495

Waer, H. (2021). The effect of integrating automated writing evaluation on EFL writing apprehension and grammatical knowledge. *Innovation in Language Learning and Teaching, 17*(1), 47-71. https://doi.org/10.1080/17501229.2021.1914062

Wang, Q., Gao, Y., & Wang, X. (2024). Exploring engagement, self-efficacy, and anxiety in large language model EFL learning: A latent profile analysis of Chinese university students. *International Journal of Human–Computer Interaction*, *41*(12), 7815-7824. https://doi.org/10.1080/10447318.2024.2400403

Wang, Y., & Fussell, S. R. (2026). "It's safer to say I don't use it": Exploring ESL students' self-perceptions as AI-assisted writers. *Proceedings of the ACM on Human-Computer Interaction*, 10(1), 1–28. https://doi.org/10.1145/3799448

Wang, Y.-Y., & Chuang, Y.-W. (2023). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, 29(4), 4785–4808. https://doi.org/10.1007/s10639-023-12015-w

Xia, S. (2026). Teacher support enhances self-efficacy and learning outcomes in the age of AI. *Scientific Reports*, 16, Article 5113. https://doi.org/10.1038/s41598-026-36014-2

Xu, S., Chen, P., & Zhang, G. (2024). Exploring the impact of the use of ChatGPT on foreign language self-efficacy among Chinese students studying abroad: The mediating role of foreign language enjoyment. *Heliyon*, *10*(21). https://doi.org/10.1016/j.heliyon.2024.e39845

Yuan, Z., & Lyu, T. (2024, June). Association between ai chatbot self-efficacy and EFL student class-related anxiety: a control-value theory perspective. In *Proceedings of the 2024 9th International Conference on Distance Education and Learning* (pp. 371-375). https://doi.org/10.1145/3675812.3675822

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), 1–37. https://doi.org/10.1186/s40561-024-00316-7

Zhang, Q., Nie, H., Fan, J., & Liu, H. (2025). Exploring the dynamics of artificial intelligence literacy on english as a foreign language learners' willingness to communicate: The critical mediating roles of artificial intelligence learning self-efficacy and classroom anxiety. *Behav Sci (Basel)*, *15*(4), 523. https://doi.org/10.3390/bs15040523

Zhao, X. (2022). Leveraging Artificial Intelligence (AI) technology for English writing: Introducing Wordtune as a digital writing assistant for EFL writers. *RELC Journal, 54*(3), 890-894. https://doi.org/10.1177/00336882221094089

Zhou, C., & Wang, Y. (2026). University Students' Writing Feedback Literacy in the AI Era: The Interplay of Generative AI Acceptance, Writing Anxiety and Writing Self-Efficacy in Chinese EMI Educational Contexts. *European Journal of Education*, *61*(1), e70478. https://doi.org/10.1111/ejed.70478

Bionote

Vo Thi Bich Phuong is currently a lecturer at HCMC University of Foreign Languages – Information Technology. She earned her MA in TESOL at HCMC University of Social Sciences and Humanities in 2016. Her academic areas of interest predominantly include learner autonomy, TESOL methodology, computer-assisted language learning, blended learning, and AI in education.

#134 INTEGRATING CHATGPT INTO EFL WRITING INSTRUCTION: EFFECTS OF THE MEDIATING ROLE OF STUDENTS' PERCEPTIONS AND PERCEIVED ACADEMIC WRITING PERFORMANCE

Nguyễn Thị Thiên Phương[1], *Nguyễn Thị Ngọc Uyển*[2]

Abstract

Generative AI tools, particularly ChatGPT, are rapidly reshaping how students approach academic writing, especially among EFL learners. Beyond providing language assistance, these technologies help reduce the pressure associated with generating ideas and structuring academic texts in English. However, the prior pieces of research have not fully clarified the mechanisms through which ChatGPT can improve students' perceived academic writing performance (PAWP), particularly the mediating role of students' perceptions in this relationship. Using data collected from 291 EFL students, this study applied PLS-SEM and bootstrapping techniques through SmartPLS4 to examine the proposed research model. The results show that the use of ChatGPT positively affects students' perceptions, which greatly enhance the degree of improvement in PAWP. Of particular interest is the direct positive effect of the tool on writing outcomes, suggesting that the effect of the tool is not merely perceptual but also directly impacts the writing process itself. These findings add to the expanding body of research on AI in education and the Technology Acceptance Model in the context of generative AI. The research also highlights the importance of increasing AI literacy in higher education and offering students explicit instructions on the ethical use of AI tools to improve academic success.

***Keywords**:* ChatGPT, academic writing performance, EFL students, perception toward AI, PLS-SEM

1. Introduction

1.1. The context of study

Higher education is undergoing a transformation due to the rapid growth of generative AI, especially in learning activities and academic writing practices. Among these technologies, ChatGPT has received much attention because of its ability to help students generate ideas, polish their language and provide instant feedback (Kasneci et al., 2023). Importantly, many universities are starting to see AI as a tool to support learning rather than simply a source of plagiarism or academic misconduct (Dwivedi et al., 2023). For EFL students this influence would seem to be even greater. In addition to critical thinking and logical argument skills, English academic writing involves good use of academic vocabulary and internationally accepted writing conventions. As a result, students increasingly use AI tools to reduce the stress of writing and to accomplish their academic duties effectively.

[1] *MA, University of Labour and Social Affairs, Campus 2, HCM City, nguyenthithienphuong78@gmail.com; phuongntt@ldxh.edu.vn*

[2] *MA, University of Labour and Social Affairs, Campus 2, HCM City, uyenntn@ldxh.edu.vn*

Nevertheless, previous studies have primarily examined the effects of ChatGPT, students' perceptions, or the writing support provided by AI in isolation. Empirical evidence simultaneously examining the relationships among ChatGPT integration, students' perceptions, and improvements in PAWP within a single mediation model remains limited, particularly in the context of EFL learners.

1.2. The rationale for the study

Although ChatGPT has been increasingly adopted, its actual effectiveness in improving PAWP remains a subject of debate. Some studies (Lo, 2023) have found that AI tools can also help students communicate ideas and produce writing more effectively. On the other hand, some scholars suggest that the over-reliance on AI may reduce the critical thinking and the academic independence of the students (Tlili et al., 2023). One of the largest gaps in the literature is the role of learners' perceptions. More specifically, limited evidence is available on how students' perceptions shape the relationship between ChatGPT use and academic writing performance. As a result, the broader educational impact of generative AI has not yet been fully understood.

1.3. Research gap

Although numerous studies have highlighted the potential of ChatGPT to support academic writing by facilitating idea generation, improving language use, and assisting with text revision (Kasneci et al., 2023; Dwivedi et al., 2023), empirical findings regarding how this technology influences learning outcomes remain inconclusive. Previous research has largely examined individual aspects of ChatGPT use, such as the extent of its adoption, students' attitudes or perceptions toward AI, or its role in supporting academic writing. Comparatively little attention has been paid to investigating the relationships among ChatGPT integration, students' perceptions toward ChatGPT, and PAWP within a single mediation framework.

Another limitation of the existing literature concerns the measurement of academic writing outcomes. Many studies rely on learners' self-reported evaluations of their writing performance without clearly distinguishing between PAWP and objectively assessed writing quality. As a result, the educational benefits of ChatGPT may be overstated if the subjective nature of these measures is not adequately acknowledged.

To address these gaps, the present study develops and tests a mediation model in which ChatGPT integration influences PAWP indirectly through students' perceptions toward ChatGPT. By examining this psychological mechanism, the study provides a more comprehensive understanding of how AI-supported learning contributes to students' writing development. It also extends the literature on technology acceptance by offering empirical evidence on the adoption of generative AI in higher education, particularly in the context of EFL academic writing.

1.4. Research objectives

This study aims to examine how the integration of ChatGPT influences PAWP of EFL students in the context of the growing use of generative AI in higher education. In addition, the research investigates the effects of students' perceptions toward ChatGPT in order to better explain how these perceptions shape the effectiveness of AI-assisted academic writing. The study therefore goes beyond evaluating the technological benefits of AI and also explores the cognitive and perceptual processes involved when students interact with AI tools in academic settings.

1.5. Research contributions

In theory, the study contributes to the literature on AI in education and expands the Technology Acceptance Model by explaining the impact of ChatGPT on the basis of students' perceptions instead of focusing on the effectiveness of the technology. The study also empirically supports a mediating model in the EFL context, helping to clarify the role of perception in shaping performance in academic writing. Methodologically, the study contributes by providing further empirical evidence from the use of PLS-SEM in higher education contexts in Asia. It thereby contributes to the growing body of research on AI-aided learning in developing economies.

2. Literature Review and hypothesis development

2.1. Literature Review

2.1.1. Integrating ChatGPT into academic settings

The development of ChatGPT is changing the way students view academic writing, especially in the EFL context. Unlike traditional grammar checkers, ChatGPT can be utilized to assist at any stage of the writing process, from brainstorming, to editing, to structuring arguments (Kasneci et al., 2023). It is a language support tool for many EFL learners and a way of reducing the stress of writing in academic English.

Another key aspect of generative AI is the opportunity for a more personalized learning experience. Students get immediate feedback and are able to adjust their writing to their needs and learning styles (Chan & Hu, 2023). Writing skills, including coherence and organization, therefore tend to improve noticeably (Zhai, 2023). Such developments indicate that ChatGPT is becoming part of the contemporary academic practices rather than merely a technical support tool.

2.1.2. Students' Perceptions of ChatGPT

In the present study, students' perceptions (PER) represent learners' overall evaluation of ChatGPT as a learning tool. Drawing on the Technology Acceptance Model (TAM), this construct encompasses perceived usefulness, perceived ease of use, attitude toward using ChatGPT, and learning satisfaction. Rather than examining these dimensions separately, the study conceptualizes them as indicators reflecting students' overall perceptions of ChatGPT in academic writing, consistent with previous AI and educational technology studies. The students' perceptions towards ChatGPT is important to how they adopt and use AI for learning activities. The construct students' perceptions toward ChatGPT (PER) based on the Technology Acceptance Model, proposed by Fred Davis (1989), is used in this study, where perceived usefulness and perceived ease of use are the main factors affecting the technology acceptance. Students who see ChatGPT as a tool that can help them to generate ideas, improve their language and support them in academic writing are more likely to develop positive attitudes towards AI. Likewise, when the tool is easy to use and provides rapid responses, students tend to engage with the technology more actively.

In addition, this study extends the original TAM framework by incorporating attitude and learning satisfaction in order to better capture students' experiences within AI-supported learning environments. These perspectives suggest that perception is not merely a personal evaluation of technology but also a cognitive mechanism that can influence the academic outcomes of EFL learners.

2.1.3. Perceived Academic Writing Performance (PAWP)

PAWP is the ability to produce a logical, accurate and consistent written work according to international academic standards. This is often a difficult skill for EFL students, as they have to juggle grammar, academic vocabulary and the organization of argument all at once. Lo (2023) found that AI-assisted writing tools can improve the quality of student writing by providing language suggestions, correcting errors, and helping students develop their content. But not all students gain equal benefit from these benefits. An important question is how students understand and use AI for academic writing.

2.2. Hypothesis Development

2.2.1. Hypothesis H1: ChatGPT Integration and Students' Perceptions

The Technology Acceptance Model suggests that users who view a technology as easy to use and useful tend to have positive attitudes toward the technology (Davis, 1989). ChatGPT can help students in academic writing to overcome challenges related to developing ideas, grammatical accuracy, and essay organization. Such advantages are expected to enhance students' perceptions of AI usefulness and usability (Sallam, 2023). Also, students' satisfaction with their learning experiences may be enhanced by instant feedback and tailored help (Chan & Hu, 2023). From these considerations the following hypothesis is proposed:

H1*: The integration of ChatGPT enhances students' perception of ChatGPT*

2.2.2. Hypothesis H2: Perceptions of Students and Academic Writing Performance

When students have positive views of ChatGPT, they are more likely to use the tool actively and effectively in the academic writing process. Such engagement can improve content organization, use of academic vocabulary, and development of arguments in English (Dempere et al., 2023).

It is important to note that positive perceptions may also improve the confidence and motivation to learn of EFL students. If students perceive AI as an accessible and helpful tool, they are more likely to be active participants in the writing process, rather than using AI as a simple replacement for doing assignments. This reasoning leads us to the following hypothesis:

H2*: Positive student perceptions toward ChatGPT have positive effects on* ***PAWP***

2.2.3. Hypothesis H3: ChatGPT Integration and Academic Writing Performance

ChatGPT may influence students' academic writing performance through the support in the writing process itself, as well as through the students' perceptions. The tool can help learners to generate ideas in a more efficient way, improve grammatical accuracy and suggest structures that are appropriate to conventions of academic writing (Kasneci et al., 2023). Immediate feedback can help improve coherence and organization of their essays as well as lessen the pressure of writing in English for EFL students. Similarly, Lo (2023) found that AI-assisted writing tools have a positive relationship with the quality of students' academic writing. Based on these arguments, the following hypothesis is formulated:

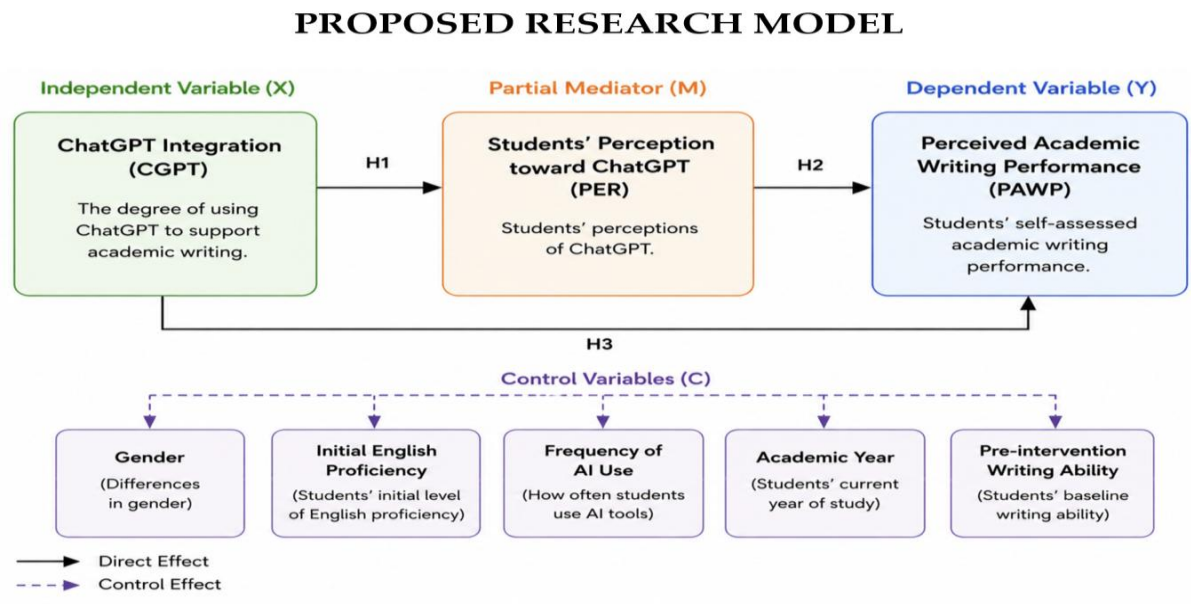
H3*: The presence of ChatGPT has a positive impact on* ***PAWP****.*

2.2.4. Hypothesis H4: The Mediating Role of Students' Perceptions

The Technology Acceptance Model states that perception is a vital mechanism to link technology use and learning outcomes. In this paper, ChatGPT is considered not only as a technical support tool, but also as a factor influencing students' assessment and acceptance of AI in academic contexts.Once positive perceptions are made, students are likely to use ChatGPT more frequently and more effectively throughout the writing process. This indicates

that perception may mediate the relationship between ChatGPT integration and academic writing performance. (Sallam, 2023). However, as AI is also expected to have a direct effect on writing performance, the mediating effect is expected to be partial rather than full. From this reasoning we propose the following hypothesis:

H4**: Students' perceptions mediate the relationship between ChatGPT Integration and **PAWP



3. Methodology

3.1. Research Design

This study used a quantitative method to explore the relationships between ChatGPT integration (CGPT), student perceptions (PER), and Perceived Academic Writing Performance (PAWP) in the context of higher education. A cross-sectional survey design was employed to obtain the perceptions and usage patterns of EFL students towards ChatGPT at a particular point of time.

The analysis was conducted using Partial Least Squares Structural Equation Modeling (PLS-SEM) due to the method's suitability for predictive research models with mediating variables (Hair et al., 2022). Another strength of PLS-SEM is that it allows for the simultaneous evaluation of the measurement model and structural model, which is particularly suitable for the investigation of learner behavior in AI-enhanced educational environments.

3.2. Research Sample and Data Collection

The participants were 291 undergraduate non-English-major students from the University of Labor and Social Affairs (Campus 2), Vietnam, who were enrolled in compulsory English courses. Data were collected by using a convenience sampling technique through an online questionnaire via Google Forms; and the survey link was distributed to students through their course instructors. Participation was voluntary, and all responses were anonymous. After screening for data analysis, a total of 291 valid responses were retained. Several procedural steps were taken to reduce the risk of common method bias. Participants were assured of the anonymity and confidentiality of their responses and informed that there were no right or wrong answers in the survey. Also, the items of the questionnaires were written in a clear and

straightforward manner in order to minimize the misunderstanding and response bias. These measures improved the reliability of self-reported data used in the study.

Several demographic and background variables, including gender, English proficiency level, frequency of AI use, and prior writing ability, were coded numerically to support additional analyses such as MGA or ANOVA when necessary.

Table 1

Sample profile

Variable Group	Category	N	Percentage (%)
Sample size (n)	Total valid responses	291	100.000
Gender	Male	128	43.986
Gender	Female	163	56.014
English Proficiency	Very low	26	8.935
English Proficiency	Low	52	17.869
English Proficiency	Intermediate	122	41.924
English Proficiency	High	62	21.306
English Proficiency	Very high	29.000	9.966
AI Use Frequency	Very infrequent	34.000	11.684
AI Use Frequency	Infrequent	52.000	17.869
AI Use Frequency	Moderate	103.000	35.395
AI Use Frequency	Frequent	68.000	23.368
AI Use Frequency	Very frequent	34.000	11.684
Academic Year	First year	78.000	26.804
Academic Year	Second year	70.000	24.055
Academic Year	Third year	77.000	26.460
Academic Year	Fourth year	66.000	22.680
Prior Writing Ability	Very low	4.000	1.375
Prior Writing Ability	Low	26.000	8.935
Prior Writing Ability	Intermediate	90.000	30.928
Prior Writing Ability	High	116.000	39.863
Prior Writing Ability	Very high	55.000	18.900

***Note**: N = 291; grouping variables are numerically coded in the dataset.*

3.3. Measurement Scales

All study variables were measured using a five-point Likert scale ranging from “strongly disagree” to “strongly agree.” The CGPT construct examined the extent to which ChatGPT assisted students in academic writing, including support for idea generation, language editing, writing completion, and ease of interaction during the writing process. Meanwhile, the PER construct was developed based on the Technology Acceptance Model (TAM), covering dimensions such as perceived usefulness, perceived ease of use, positive attitude, and learning satisfaction.

PAWP was measured by several aspects of writing quality such as accuracy of grammar, coherence and cohesion, vocabulary use, writing organization and Overall writing improvement. These measurement scales not only captured students' ability to use technology effectively but also reflected how they engaged with academic writing in an EFL learning environment.

The outer loading values for all indicators exceeded the recommended threshold of 0.70, indicating satisfactory indicator reliability within the PLS-SEM model.

Table 2

Measurement items and outer loadings

Construct	Item	Measurement Content (Shortened)	Mean	SD	Outer loading	Conclusion
CGPT	CGPT1	Idea generation support	3.137	1.260	0.856	Retained
CGPT	CGPT2	Language editing support	3.007	1.268	0.861	Retained
CGPT	CGPT3	Writing completion support	3.189	1.279	0.879	Retained
CGPT	CGPT4	Ease of interaction	3.134	1.267	0.858	Retained
PER	PER1	Perceived usefulness	3.052	1.265	0.855	Retained
PER	PER2	Perceived ease of use	3.131	1.222	0.866	Retained
PER	PER3	Positive attitude	2.959	1.231	0.843	Retained
PER	PER4	Learning satisfaction	2.952	1.228	0.853	Retained
PAWP	PAWP 1	Grammar accuracy	3.052	1.257	0.862	Retained
PAWP	PAWP 2	Coherence and cohesion	2.918	1.295	0.879	Retained
PAWP	PAWP 3	Vocabulary use	3.003	1.222	0.857	Retained
PAWP	PAWP 4	Writing organization	3.076	1.274	0.848	Retained
PAWP	PAWP 5	Overall writing improvement	2.997	1.250	0.829	Retained

***Note**: All outer loadings exceed the 0.70 threshold.*

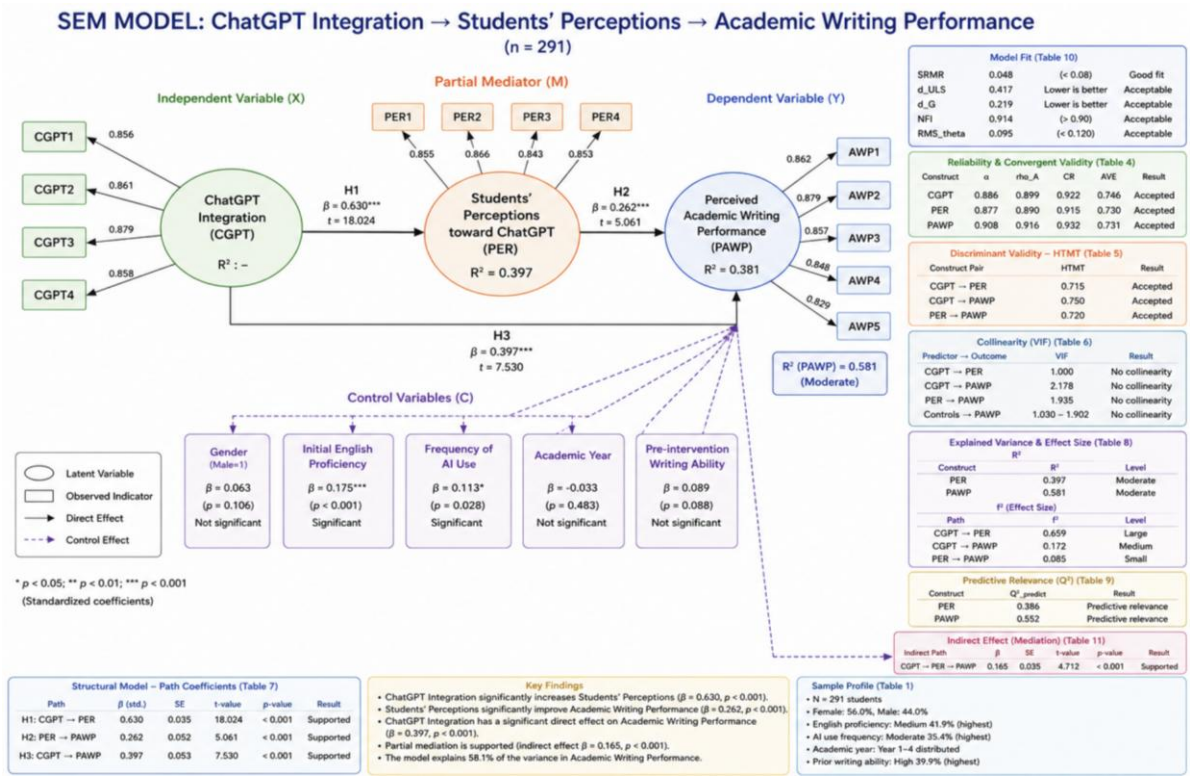
3.4. Data analysis method

The data were analyzed using SmartPLS 4 with the Partial Least Squares Structural Equation Modeling (PLS-SEM) approach. The evaluation process consisted of two main stages: assessment of the measurement model and assessment of the structural model.

For the measurement model, several indicators were examined, including outer loadings, Cronbach's alpha, Composite Reliability, Average Variance Extracted (AVE), and the Heterotrait-Monotrait ratio (HTMT), in order to evaluate the reliability and discriminant validity of the constructs.

In the structural model assessment, hypothesis testing was conducted through bootstrapping with 5,000 resamples. Model fit, explanatory power and predictive relevance were assessed using indices such as SRMR, R2, f2 and Q2. In addition, Variance Inflation Factor (VIF) values were assessed to reduce the possibility of multicollinearity among the research variables.

4. Results



4.1. Mmeasurement model Assessment

The measurement model was evaluated in terms of reliability, convergent validity, discriminant validity, and multicollinearity following the guidelines proposed by Hair et al. (2022). The results showed that all the constructs met the recommended standards. Cronbach's Alpha and Composite Reliability were greater than 0.70 and all AVE values were above the cut-off value of 0.50.

The results revealed that the measurement scales possessed good reliability and strong convergent validity. PAWP has the highest Composite Reliability value (0.932) among the constructs which indicates the high level of internal consistency among the observed indicators of PAWP.

To assess the possible effects of common method bias, a Harman's single-factor test was carried out. The largest factor accounted for 38.4% of the total variance which was lower than the recommended cutoff of 50%, suggesting that common method bias was less likely to be a serious concern in this study. Moreover, all the complete collinearity VIF values were below 3.30, which provides additional evidence for the absence of significant bias due to self-reported data.

Furthermore, all HTMT values were lower than 0.85, confirming satisfactory discriminant validity and indicating that the constructs were conceptually distinct from one another.

Table 3

Reliability and convergent validity

Construct	Number of Items	Cronbach's Alpha	rho_A	CR	AVE	Conclusion
CGPT	4.000	0.886	0.899	0.922	0.746	Accepted

PER	4.000	0.877	0.890	0.915	0.730	Accepted
PAWP	5.000	0.908	0.916	0.932	0.731	Accepted

***Note**: Cronbach's Alpha and CR exceed 0.70; AVE exceeds 0.50.*

In addition, discriminant validity was evaluated using the Heterotrait-Monotrait Ratio (HTMT) of correlations.. As presented in Table 4, all HTMT values were below the recommended threshold of 0.85, confirming that the constructs in the model were clearly distinguishable from one another. These findings indicate that the variables represented separate concepts and that no significant overlap among the constructs was observed.

Table 4

Discriminant validity

Construct Pair	Comparison	HTMT	Conclusion
CGPT	PER	0.715	Accepted
CGPT	PAWP	0.750	Accepted
PER	PAWP	0.720	Accepted

***Note**: HTMT values are below 0.85 indicating discriminant validity.*

To check for any multicollinearity between the constructs, Variance Inflation Factor (VIF) values were tested. From the VIF results in Table 5, all values were less than the threshold of 3.30, indicating that there was no serious multicollinearity issue in the model. The finding supports the stability and reliability for the regression coefficients of the structural model.

Table 5

Collinearity assessment

Predictor Variable	Dependent Variable	VIF	Conclusion
CGPT	PER	1.000	No collinearity
CGPT	PAWP	2.178	No collinearity
PER	PAWP	1.935	No collinearity
Gender	PAWP	1.030	No collinearity
English Proficiency	PAWP	1.552	No collinearity
AI Use Frequency	PAWP	1.902	No collinearity
Academic Year	PAWP	1.254	No collinearity
Prior Writing Ability	PAWP	1.599	No collinearity

***Note**: VIF values are below 3.30, indicating no serious collinearity issue.*

4.2. Structural Model Assessment

The structural model was examined using the bootstrapping procedure with 5,000 resamples in SmartPLS 4. The findings revealed that all three main hypotheses were supported at the significance level of $p < 0.001$. Specifically, the integrated ChatGPT had a positive impact on students' perceptions ($\beta = 0.630$). This means that the use of AI enhanced students' perceived usefulness and perceived ease of use. Furthermore, students' perceptions had a positive effect on PAWP ($\beta = 0.262$). This finding highlights the important role of perception in improving academic writing outcomes.

Apart from the indirect effect, the integration of ChatGPT also had a direct positive effect on PAWP (β = 0.397). This indicates that AI tools can directly help students in generating ideas, improving language, and better organization of written content. Among the control variables, only English proficiency and frequency of AI use had statistically significant effects on PAWP.

The results of R^2 showed that the model has a moderate explanatory power of 0.397 for PER and 0.581 for PAWP. Furthermore, the path from CGPT to PER produced a large effect size (f^2 = 0.659). All Q^2 values were greater than zero, confirming that the model possessed satisfactory predictive relevance for the endogenous constructs.

Table 6

Structural model results

Path	β	SE	t-value	p-value	LLCI	ULCI	Conclusion
H1: CGPT → PER	0.630	0.035	18.024	<0.001	0.558	0.694	Supported
H2: PER → PAWP	0.262	0.052	5.061	<0.001	0.162	0.362	Supported
H3: CGPT → PAWP	0.397	0.053	7.530	<0.001	0.294	0.497	Supported
Gender → PAWP	0.063	0.039	1.617	0.106	-0.012	0.138	Not significant
English proficiency → PAWP	0.175	0.048	3.662	<0.001	0.083	0.268	Significant
Frequency of AI use → PAWP	0.113	0.051	2.197	0.028	0.013	0.214	Significant
Academic year → PAWP	-0.033	0.047	-0.701	0.483	-0.125	0.060	Not significant
Prior writing ability → PAWP	0.089	0.052	1.704	0.088	-0.012	0.192	Not significant

***Note**: Bootstrapping with 5,000 subsamples; standardized coefficients are reported.*

The results indicated that the integration of the AI system had a direct positive effect on PAWP (β = 0.397), suggesting the supportive role of AI in idea generation, language editing, and the organization of academic writing. Among control variables, only English proficiency and AI use frequency were statistically significant predictors of PAWP.

Moreover, the model showed moderate explanatory power with R2 of 0.397 for PER and 0.581 for PAWP. CGPT and PER were significantly related to each other with a large effect size (f^2 = 0.659). The effect of PER on PAWP was relatively weak.

Table 7

Coefficient of Determination and Effect Size (R^2, f^2)

Indicator	Variable / Path	Value	Level
R^2	PER	0.397	Moderate
R^2	PAWP	0.581	Moderate
f^2	CGPT → PER	0.659	Large
f^2	CGPT → PAWP	0.172	Medium
f^2	PER → PAWP	0.085	Small

f^2	Gender → PAWP	0.009	Very small
f^2	English_Proficiency → PAWP	0.047	Small
f^2	AI_Use_Frequency → PAWP	0.016	Very small
f^2	Academic_Year → PAWP	0.002	Very small
f^2	Prior_Writing_Ability → PAWP	0.012	Very small

***Note**: r2 = explanatory power f2 = predictor level effect size.*

Regarding model fit, the SRMR value is 0.048 (Table 8) which is below the recommended threshold of 0.08. The model fits well the observed data.

Table 8

Model Fit Indices (SRMR)

Index	Value	Reference Threshold	Conclusion
SRMR	0.048	<0.080	Good fit
d_ULS	0.417	Lower is better	Acceptable
d_G	0.219	Lower is better	Acceptable
NFI	0.914	>0.900	Acceptable
RMS_theta	0.095	<0.120	Acceptable

***Note**: SRMR below 0.08 indicates acceptable model fit.*

The results of indirect effect in Table 9 indicate that PER partially mediates the relationship between CGPT and PAWP, as the indirect effect was statistically significant (β = 0.165; p <0.001). In addition, the bootstrap confidence interval did not include zero, thereby confirming the role of cognitive perceptions in translating AI integration into improved PAWP.

Table 9

Indirect effects / mediation results

Indirect effects	β	SE	t-value	p-value	LLCI	ULCI	Conclusion
CGPT → PER → PAWP	0.165	0.035	4.712	<0.001	0.099	0.237	Partial mediation supported

***Note**: The bootstrap confidence interval excludes zero, supporting partial mediation.*

5. Discussion

5.1. The Effect of CGPT on PER

The results indicate that CGPT has a positive effect on PER, suggesting that students' experiences with the chat-based generative AI have changed EFL learners' perceptions regarding the role of AI in the academic setting. An aspect worth mentioning is the chatbot's capacity to give almost instant feedback on grammar, vocabulary and essay organization, which may enhance students' perceptions of the technology as both practical and user-friendly.

In addition, ChatGPT seems to function as a type of cognitive scaffolding, guiding students through the various stages of the academic writing process. Such support can help reduce cognitive load, so that students can devote more attention to argument construction and improving coherence in their writing. These results are in line with the study by Kasneci et al. (2023) who argued that generative AI has strong potential to deliver personalized learning support within higher education environments

5.2. The direct effect of CGPT on PAWP

The findings revealed that CGPT has a direct influence on PAWP. This suggests that ChatGPT contributes to EFL students' academic writing not only through their perceptions of the technology but also through direct support during the writing process itself.

One possible reason is its reduced cognitive load for students when they write in English. ChatGPT helps learners to focus more on idea generation and academic reasoning by assisting with language correction and providing content-related suggestions. In the same way, the tool promotes self-regulated learning, as learners can autonomously revise, edit and improve their writing thanks to the immediate feedback. As a result, students are able to improve coherence and cohesion, essay organization, and use of academic vocabulary in their writing assignments.

5.3. The partial mediating role of PER

The results suggest that PER partially mediates the relationship between CGPT and PAWP. This suggests that the success of ChatGPT is not necessarily due to the technical capabilities of AI but rather to students' perceptions and interactions with the technology in an academic setting. When learners have a positive perception of the tool, they are more likely to use the tool actively and effectively during the process of academic writing.

However, the mediating role of PER was not complete but partial, indicating that ChatGPT still has a direct effect on PAWP. In other words, AI seems to be providing cognitive support and direct assistance in academic tasks (e.g., grammar correction, paper structuring, academic expression improvement).

6. Theoretical implications of the study

This study contributes to the growing AI-in-education literature by clarifying the mechanism through which ChatGPT influences PAWP via students' perceptions. Unlike many earlier studies that mainly emphasized the technical or ethical dimensions of AI, the present findings suggest that the effectiveness of generative AI also depends on how learners perceive, adopt, and interact with the technology in academic environments.

From a theoretical perspective, the study extends the TAM within the context of generative AI. The findings indicated that perceived usefulness and perceived ease of use are associated not only with technology acceptance but also with academic outcomes, particularly PAWP. In addition, PER is conceptualized as a higher-order construct consisting of usefulness, ease of use, attitude, and satisfaction, allowing for a more comprehensive understanding of students' learning experiences in AI-supported educational settings.

The partial mediating role of PER also reveals that AI has more than a cognitive support; it could also have a direct effect on learning processes and the achievement of academic tasks. The findings offer further empirical evidence for research on EFL learning and AI-supported education in the context of higher education in Asia.

7. Practical implications

The results of the PLS-SEM indicate that ChatGPT has a significant impact on students' perceptions and PAWP. The results indicate that universities should not adopt a prohibitive attitude towards AI but should include AI literacy in education programs. Students need to be trained in how to write effective prompts, critically evaluate academic sources, and understand the limitations of AI in order to use the technology in a more informed and reflective way. This is in accordance with the arguments presented by Kasneci et al. (2023)

Additionally, as PER only partially mediated the relationship between ChatGPT and PAWP, the results suggest that ChatGPT affects students' writing not only by their perception of the technology but also by providing direct assistance in academic tasks. This is why educational institutions need to set clear guidelines for ethical use of AI to reduce risks associated with plagiarism and over-reliance on AI-generated content. As Dwivedi et al. argue, it is important for students to be aware of the difference between using AI as a support tool and AI as a total replacement of independent academic thinking.

The findings also indicate that although ChatGPT should be regarded as a writing assistant, it should not be regarded as a replacement for academic writing per se. Faculty can encourage critical engagement by requiring students to submit early drafts, AI-generated feedback, and reflective self-assessments as part of the writing process. These practices can be useful for strengthening critical thinking, and for supporting the development of academic writing skills.

8. Limitations and future research directions

The study has made a contribution but has a number of limitations that should be considered in interpreting the findings. Firstly, the cross-sectional design describes the relationship between variables at one point in time, this limits the ability to show clear causal relationships between. Second, the data were collected through self-report surveys, and, therefore, the results may be influenced by personal perceptions and response bias. Specifically, the PAWP variable mainly reflects students' subjective evaluations rather than actual writing scores or instructor ratings, and therefore may not fully capture learners' authentic academic writing ability. Although Harman's single-factor test and the full collinearity VIF assessment suggested that common method bias was not a serious concern, these statistical procedures cannot completely eliminate the possibility of common method variance inherent in self-reported cross-sectional data.

Moreover, convenience sampling restricts the generalizability of the findings to other EFL contexts. Because the participants were recruited from a single university in Vietnam, the findings should be interpreted with caution when generalizing to students from other institutions or educational settings. In addition, the use of voluntary participation may have introduced self-selection bias, as students who were more interested in or familiar with ChatGPT might have been more willing to participate in the survey. The present study was not able to fully capture differences between groups of students in academic culture, digital competence, and the level of access to AI technologies. Furthermore, although the questionnaire items were adapted from validated instruments, the absence of formal pilot testing and/or expert review should also be acknowledged as a methodological limitation, as additional evidence of content validity could further strengthen the measurement instrument.

Future research can overcome these limitations by adopting longitudinal designs to study the evolution of AI usage behaviors over time. The inclusion of instructor evaluations or standardized writing rubrics would also be useful in providing a more objective measure of academic writing performance. Furthermore, techniques such as measurement invariance assessment and MGA in PLS-SEM and variables like AI literacy and technology readiness might help in improving the explanatory power of future research models.

9. Conclusion

The study demonstrates that ChatGPT has a significant influence on PAWP through both direct effects and indirect effects mediated by learners' perceptions. This indicates that the success of AI in education depends not only on technological capabilities but also on how

students perceive, evaluate and use AI tools in the learning process. When students develop positive perceptions of ChatGPT, they are more likely to use the tool proactively to generate ideas, refine language use, and improve the organization of academic writing.

The findings also indicate that ChatGPT serves not only as a technical support tool but also as a factor that can directly influence students' learning and academic writing processes. In this sense, generative AI may help reduce the pressure associated with writing in English while simultaneously encouraging greater self-regulation and the development of academic thinking skills.

Theoretically, the study extends the TAM in the context of AI in education and adds more empirical evidence to the EFL literature, especially in higher education settings in developing economies. Practical implications The findings demonstrate that the appropriate integration and management of ChatGPT as a writing assistant can be effective in university learning environments. Hence, higher education institutions should pay more attention to AI literacy and ethical use of AI to help students use AI as a learning support tool instead of over-reliance on the technology.

References

Chan, C. K. Y., & Hu, W. (2023). *Students' voices on generative AI: Perceptions, benefits, and challenges in higher education*. International Journal of Educational Technology in Higher Education, 20, 43. https://doi.org/10.1186/s41239-023-00411-8

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M., Al-Busaidi, K. A., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., & Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage Publications.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare, 11*(6), Article 887. https://doi.org/10.3390/healthcare11060887

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). *The impact of ChatGPT on higher education*. Frontiers in Education, 8, 1206936. https://doi.org/10.3389/feduc.2023.1206936

Zhai, X. (2023). ChatGPT for next generation science learning. *XRDS: Crossroads, The ACM Magazine for Students, 30*(2), 42–46. https://doi.org/10.1145/3624720

Appendix

A Five-Point Likert Scale Used for the Survey Questionnaire

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Code	Scale item	1	2	3	4	5	Notes
CGPT1	I use ChatGPT to generate ideas when beginning an academic writing task in English.	☐	☐	☐	☐	☐	Adapted from theories of technology acceptance and digital learning support: Davis (1989); Venkatesh et al. (2003); Kasneci et al. (2023).
CGPT2	I use ChatGPT to improve grammar, vocabulary, and sentence expression in my English writing.	☐	☐	☐	☐	☐	Adapted from theories of technology acceptance and digital learning support: Davis (1989); Venkatesh et al. (2003); Kasneci et al. (2023).
CGPT3	I use ChatGPT to develop outlines, paragraphs, and the overall structure of academic essays.	☐	☐	☐	☐	☐	Adapted from theories of technology acceptance and digital learning support: Davis (1989); Venkatesh et al. (2003); Kasneci et al. (2023).
CGPT4	I find interacting with ChatGPT during the writing process convenient and easy to do.	☐	☐	☐	☐	☐	Adapted from theories of technology acceptance and digital learning support: Davis (1989); Venkatesh et al. (2003); Kasneci et al. (2023).

Code	Scale item	1	2	3	4	5	Notes
PER1	I believe that ChatGPT is beneficial for improving the process of learning academic writing in English.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER2	I find ChatGPT easy to use when seeking support for academic writing tasks.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use,

							attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER3	I have a positive attitude toward using ChatGPT in English writing classes.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER4	I am satisfied with the level of support that ChatGPT provides during academic writing activities.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).

Code	Scale item	1	2	3	4	5	Notes
AWP1	After the learning process, my English writing shows greater grammatical accuracy.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP2	After the learning process, my writing has become more coherent and better connected.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP3	After the learning process, I use academic vocabulary more appropriately in my writing.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP4	After the learning process, I am able to organize ideas and structure essays more clearly.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP5	Overall, my English academic writing performance has improved.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003);

							Weigle (2002); Graham & Perin (2007).

Code	Scale item	1	2	3	4	5	Notes
PER1	I believe that ChatGPT is beneficial for improving the process of learning academic writing in English.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER2	I find ChatGPT easy to use when seeking support for academic writing tasks.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER3	I have a positive attitude toward using ChatGPT in English writing classes.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).
PER4	I am satisfied with the level of support that ChatGPT provides during academic writing activities.	☐	☐	☐	☐	☐	Adapted from constructs related to perceived usefulness, ease of use, attitudes, and learning satisfaction: Davis (1989); Venkatesh et al. (2003); Teo (2011).

Code	Scale item	1	2	3	4	5	Notes
AWP1	After the learning process, my English writing shows greater grammatical accuracy.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP2	After the learning process, my writing has become more coherent and better connected.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003);

							Weigle (2002); Graham & Perin (2007).
AWP3	After the learning process, I use academic vocabulary more appropriately in my writing.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP4	After the learning process, I am able to organize ideas and structure essays more clearly.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).
AWP5	Overall, my English academic writing performance has improved.	☐	☐	☐	☐	☐	Adapted from academic writing assessment criteria: Hyland (2003); Weigle (2002); Graham & Perin (2007).

#135 AN ANALYSIS ON EFL VIETNAMESE SOPHOMORES PRONUNCIATION IN A SPEAKING COURSE

Tran Thi Minh Thu[1]

Abstract

This study aims to investigate the pronunciation features of EFL second year university students in a speaking course, focusing on consonant production and prosodic elements such as word stress and intonation. To demonstrate the significant difficulties that learners encountered during producing English speech, the two participants were asked to perform both read-a-loud tasks and speaking tasks. Data from recorded speech were analyzed base on the phonological frameworks of previous research (Roach, 2009; Rogerson-Revell, 2011; Cruttenden, 2014). The findings reveal that the participants have the tendency of omitting the final sounds such as -s and -ed, and have troubles in producing a good English intonation, affecting their communicative effectiveness. Consistent with prior research (Derwing & Munro, 2005; Yates, 2017; Pourhosein, 2016), this small qualitative case study underscores the importance of targeted pronunciation instruction, calling for further research to implement pedagogical interventions which emphasize prosodic features and final consonants to enhance students' overall spoken proficiency.

[1] *Tran Thi Minh Thu thuttm@huflit.edu.vn Faculty of Foreign Languages Ho Chi Minh City University of Foreign Languages and Information Technology*

***Keywords*:** Vietnamese EFL learners, pronunciation, consonants, word stress, intonation

1. Introduction

Pronunciation has caught a lot of attention of many researchers when it comes to learning English. Noticeably, L2 learners have to put a lot of efforts to improve their pronunciation as it is a difficult part of English learning process (Pourhosein, 2016). Besides, it is claimed that pronunciation plays a crucial part in supporting learners' understanding through spoken communication (Derwing & Munro, 2005; Yates, 2017). Thus, the importance of English pronunciation is taken into account in both learning and teaching process. In relations to pronunciation features, stress and consonants are supposed to play important parts in a comprehension conversation. Noticeably, Vietnamese students usually made errors in these aspects of pronunciation, which causes the difficult understanding in communicating with native speakers.

Previous studies have consistently reported that Vietnamese learners of English experience difficulties with English final consonants and consonant clusters due to substantial phonological differences between Vietnamese and English (Nguyen & Newton, 2020; Hien, 2024; Bui et al., 2021). Research has also shown that suprasegmental features such as word stress and intonation remain challenging for Vietnamese learners because Vietnamese is a tonal language and does not employ stress and intonation in the same manner as English (Yates, 2017). Although these pronunciation difficulties have been documented, much of the existing research has focused on larger groups of learners and quantitative descriptions of error types.

Sophomore English majors represent an important population for investigation because they have already received substantial English instruction and are expected to participate in increasingly demanding speaking courses. At this stage, persistent pronunciation difficulties may affect their future academic performance, professional communication, and language development. Furthermore, examining pronunciation within a speaking-course context has pedagogical value because the findings can directly inform classroom pronunciation instruction. Hence, this paper aims to explore these aspects related to Vietnamese learners' pronunciation and some appropriate recommendations for pronunciation teaching in Vietnam.

2. Literature review

The paper analysis will be based on the pronunciation features stated below:

2.1. Consonants

In terms of English pronunciation, Rogerson-Revell (2011) suggested that there are 24 consonants described as the table below:

Table 1

English consonants phonemes

Place of articulation / Manner of articulation	Bilabial	labiodental	dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
PLOSIVE	p b			t d			k g	
FRICATIVE		f v	θð	s z	ʃʒ			h
AFFRICATE					ʧ ʤ			
NASAL	m			n			ŋ	

It can be seen that the consonants are classified into manner and place of articulation. Additionally, these consonants are divided into voiced group (/b/, /m/, /w/, /v/, /d/, /z/, /n/, /l/, /r/, /j/, /g/, /ʒ/, /dʒ/, /ŋ/, /ð/) and voiceless group (/p/, /f/, /t/, /s/, /k/, /h/, /ʃ/, /tʃ/, /θ/). Noticeably, the consonants made up a large proportion for ending sounds in English, especially for s rule pronunciation and past tense verbs pronunciation. These ending sounds are usually known as codas (final consonants) which is a crucial part in the syllable structure (Roach, 2009). Moreover, consonants clusters happen when there are two or more consonants combined together (Roach, 2009). In comparison to Vietnamese, most of the consonants share the same sounds (see Appendix A), so it is likely that Vietnamese can produce consonant clusters accurately. However, it is claimed that L2 learners usually struggle with pronouncing the final sounds relating to English rules. Thus, it is necessary to carry out the errors and give feedback to learners for further improvement in their pronunciation.

2.2. Final consonant articulation among Vietnamese learners

Regarding pronunciation production, this has been proved that second language learners were affected by their mother tongue to some extent. Although Vietnamese and English shared some consonants such as /b, d, k, m, n, f, v, s, z, h, l/ (Pham & McLeod, 2016), Vietnamese does not have the habit of pronouncing final sounds, leading to the ignorance of final sounds or difficulties articulating these sounds among Vietnamese second language learners (Nguyen & Newton, 2020). As mentioned above, consonants are divided into voiceless and voiced consonants, along with consonant clusters, which poses challenges for non-native learners (Al-khresheh, 2024). Following the consonants categories described as above (Rogerson-Revell, 2011), Cummins (2015) stated that most English words ended with these consonants, except for except for /h/, /r/, /w/ and /j/. In addition to this, Knight (2012) listed final consonant sounds as below:

/p/ - sheep, stop, cheap

/b/ - transcribe, crab

/t/ - start, smart, seat, meat, wet

/d/ - ride, sad, bread, road,

/k/ - pick, speak, peak

/f/ - leaf, belief, roof, half, sniff

/v/ - leave, arrive, live, love, five

/θ/ - earth, health, worth

/ð/ - breathe, clothe, with

/s/ - stress, goes, rice, bus, six

/z/ - these, plays, buzz, prize

/tʃ/ - teach, watch, much

/dʒ/- bridge, large, village

/m/- come, some, warm, time

/n/ - than, man, sun, tin, cone

/ŋ/ - sing, spring, wrong, wing

/l/ - pool, smile, veil, call, girl

/ʒ/ - message, garage, massage

/g/ - beg, dog, clog, bag

/ʃ/ - crash, wash, rush

In addition to this list, the researcher categorized final consonants into two groups, namely pre-final and post-final consonants.

Pre-final includes: /m/, /n/, /ŋ/, /l/, /s/

Post-final includes: /s/, /z/, /t/, /d/, /θ/

The study of Bui et al. (2021) examined how Vietnamese learners pronounced these consonants stated above. The researchers found out that most learners mispronounced these sounds or omitted them when encountering these consonants during their productive speech. In some cases, learners also substituted these sounds, leading to the misunderstanding or incorrectness to listeners. As a result, these are common errors in pronouncing ending sounds among Vietnamese second language learners.

Another study from Hien (2024) showed that Vietnamese learners had difficulties pronouncing ending sounds of English words due to the differences between Vietnamese and English consonants systems. The study collected data from five participants representing different Vietnamese regional accents. Participants were asked to read aloud a paragraph provided by the researcher and submit their recordings via email. The findings revealed that learners frequently omitted unfamiliar or challenging final consonants and inserted a schwa vowel into final consonant clusters. For example, lives was pronounced as /lɪvəz/ and leaves as /liːvəz/. The participants also tended to simplify final consonants by modifying them to resemble sounds found in their native language. Additionally, the quality of two recordings was compromised by background noise, making accurate phonetic analysis difficult.

Although there are various pronunciation topics, many researchers seem to disregard these common errors in pronouncing final consonants among Vietnamese second language learners. Specifically, students who majored in English Language also encountered these errors in their pronunciation, leading to the need of further research in this aspect and appropriate pedagogical practices for improvements in this field.

2.2. Word Stress and Intonation

2.2.1. Word stress

Cruttenden (2014) referred the concept of stress as ‘the syllable or syllables of a word which stand out from the remainder are said to be accented, to receive the accent’. Additionally,

there are four main factors that is made to recap syllable prominence or stress (Rogerson-Revell, 2011):

1. Loudness
2. Pitch change
3. Quality of sound
4. Length of sound

Previous research suggests that Vietnamese learners often demonstrate reduced pitch variation, inaccurate stress placement, and difficulties with English prosodic features because English stress patterns differ substantially from Vietnamese phonological organization (Yates, 2017; Rogerson-Revell, 2011). It is considered that pitch change is the most important part to recognize the stressed and unstressed syllables. Hence, learners are supposed to have native-like voice if they correctly process these four parts in making word stress.

2.2.2. Intonation

In terms of broader aspects of stress, Rogerson-Revell (2011) defined this as prosody or speech melody which builds up English intonation. To specify English intonation, Rogerson-Revell (2011) provided the main structure as below:

- Thought groups/Tone units (the division of utterances into smaller chunks or meaningful units)
- Nuclear stress placement/Tonicity (the accentual function of intonation)
- Tone choice (Rise-Fall/pitch movement)

As Roach (2009) suggested, English intonation enables learners to engage with attitudinal, accentual, grammatical, discourse function which support their second language uptake. Therefore, it is important for learners to take into account these aspects of pronunciation for improving their articulation. Hence, this paper will apply these features to make a clear examination of Vietnamese learners' pronunciation.

2.4. Research questions:

Collectively, previous studies suggest that final consonants, word stress, and intonation remain problematic for Vietnamese learners of English. However, existing research has largely focused on identifying error patterns in broader learner populations. Few studies have provided an in-depth qualitative examination of how these pronunciation features appear across both controlled and spontaneous speaking tasks among university English majors. Therefore, the present study addresses this gap through a small-scale case analysis of two Vietnamese EFL sophomores.

a. What consonant difficulties do Vietnamese EFL sophomores encounter during both their read-a-loud and free speaking task?

b. What are some common mistakes do Vietnamese EFL sophomores make regarding English word stress and intonation?

c. What are the possible implications of these pronunciation difficulties for communicative effectiveness?

3. Methodology

3.1. Research design and procedure

This study employed a qualitative descriptive case-study design to examine the pronunciation features and difficulties of EFL sophomores in a speaking course at a university. The purpose of the study was not to generalize findings to all Vietnamese EFL learners but to provide an in-depth examination of pronunciation features exhibited by two information-rich cases. A case-study approach was considered appropriate because it allows detailed analysis of naturally occurring pronunciation patterns across different speaking tasks and generates pedagogical insights that may inform pronunciation instruction.

The participants are asked to record their pronunciation for this paper analysis. The speakers are given an English passage to read, and they are required to give a free speech. Before the process of recording, the participants have filled in the consent form and questionnaire to give their background information. The analysis is carried out based on the data given by the learners.

3.2. Participants

There were two Vietnamese EFL sophomores taking part in this qualitative case study. The participants were chosen based on their on-going assessment results from their lecturer within their classroom. Purposive sampling was employed to select two information-rich participants for detailed analysis. To ensure a representative cross-section for most Vietnamese EFL sophomores, the two learners with the same proficiency level (B2) were identified as fairly good pronunciation and good pronunciation according to their lecturer. The participants' native language is Vietnamese, and they had studied English as a compulsory subject through Vietnamese education system for approximately ten years. The two participants are majoring in English language, and they provided rich qualitative data that allowed for an in-depth examination of pronunciation features commonly reported in previous studies on Vietnamese learners of English. The purpose of this selection was not to represent the broader population of Vietnamese EFL sophomores but to examine persistent pronunciation difficulties among learners who had already achieved relatively high levels of English proficiency.

Because both participants shared similar educational backgrounds and proficiency levels, the findings were interpreted as exploratory rather than representative. It is possible that learners with different proficiency levels or learning experiences may demonstrate different pronunciation patterns.

3.3. Data collection instruments

A free speaking task (see Appendix B) and reading passages adopted from (Fairbanks, 1960; Darley et al., 1975) were deployed as primary instruments to assess learners' pronunciation. The two reading passages were selected since it contains numerous final consonants, consonant clusters, plural markers, and multisyllabic words, making it suitable for examining consonant production and word stress patterns. An oral test acted as free speaking task to obtain the data of phonetic nature for the analysis of targeted sounds. The free speaking task had only one part with 4 questions following the familiar topic that learners studied during their speaking course (described in appendix B). This task allows learners to perform spontaneous speech and manage their language production, supporting this study obtain a more comprehensive understanding of their pronunciation strengths, weaknesses, and learning goals. Read-aloud passages adapted from a pronunciation assessment text were used to elicit controlled speech. The passage contained a variety of consonant sounds, final consonants, consonant clusters, multisyllabic words, and opportunities for intonation patterns. This task enabled the researcher to identify pronunciation difficulties under relatively controlled conditions.

3.4. Data analysis

All recordings were transcribed and repeatedly listened to for detailed phonological analysis. This study examined the two learners' pronunciation based on the framework proposed by Roach (2009), Rogerson-Revell (2011), and Cruttenden (2014) as mentioned in the literature review. The analysis mainly focused on final consonants articulation, word stress, and intonation. Particular attention was paid to the pronunciation of plural and third-person singular markers (-s/-es) and past tense markers (-ed), as these structures frequently pose difficulties for Vietnamese learners (Disney & Le, 2024).

To improve analytical reliability, all recordings were transcribed and reviewed multiple times using a consistent coding sheet focusing on final consonants, word stress, and intonation. Pronunciation errors were identified according to the phonological descriptions provided by Roach (2009), Rogerson-Revell (2011), and Cruttenden (2014). Because the study was exploratory and small-scale, formal inter-rater reliability was not conducted. This limitation is acknowledged and should be addressed in future research.

Regarding validity, the combination of read-aloud and free-speaking tasks enabled the study to examine pronunciation features under both controlled and spontaneous speaking conditions. The selected tasks were considered appropriate because they provided opportunities for the production of final consonants, multisyllabic words, stress patterns, and intonation structures relevant to the research questions.

4. Analysis and discussion

4.1. Consonants

As can be heard from the recordings, all the participants can pronounce /s/, /h/, /g/, /f/, /b/, /n/, /m/, /r/, /w/ the same as native speakers. In comparison to L1, these sounds share the same sound in Vietnamese (see Appendix A), so it is obvious that learners can articulate these sounds correctly. With regard to the consonant clusters, the participants seemed to have a good performance of pronouncing them. However, during reading a passage aloud, the participants did make errors for some consonants (as described in table 1), namely /z/, /t/, /tʃ/. Interestingly, these participants are supposed to mostly pronounce incorrectly these sounds since their proficiency is B2.

Table 2

Errors in pronouncing consonants sounds (reading aloud passage)

Consonants	Errors
/z/	Instead of pronouncing [z] for the words observe /əbˈ**z**ɜːv/ and ooze /uː**z**/, speaker 1 pronounced [s] for these words: /əbˈ**s**ɜːv/, /uː**s**/ (RP Speaker 1, 0:22)
/t/	The speaker 1 did not pronounce it clearly for the word twice in the passage, she got a half sound of /t/ and a half sound of pronouncing 'ch' in Vietnamese. (RP Speaker 1, 0:34)
/tʃ/	The speaker 2 mispronounced the word arch /ɑːr**tʃ**/, she said /eɪ**k**/ instead. (RP Speaker 2, 0:15)

Noticeably, Vietnamese learners seem to fail in pronounce the final codas correctly, especially for the s rules pronunciation. As can be heard from the recordings, the learners pronounce /s/ for all of the words ending with s. In both reading aloud the passage and free speech, learners made no differences of final consonant sounds (/z/, /s/, /ɪz/) for words belonging to the s rules pronunciation. Interestingly, learners also did not produce correctly /t/, /d/ for the verbs in -ed form. The most salient feature can be found that /t/ was mostly unaspirated for words ending with 't' as well as words engaging with -ed rule.

Table 3

Final consonants errors

Feature	Target tokens	Errors	Error rates
/z/	15	10	66.7%
/ɪz/	8	5	62.5%
/d/	12	7	58.3%
/t/	14	6	42.9%

Table 3 demonstrates that the greatest pronunciation difficulties occurred in the production of English morphophonemic final consonants, particularly /z/ (66.7%) and /ɪz/ (62.5%), followed by the past tense marker /d/ (58.3%) and the voiceless marker /t/ (42.9%). These findings indicate that the learners' errors were not randomly distributed across English consonants but were concentrated on grammatical morphemes realized through word-final consonants. This pattern supports the phonological framework proposed by Roach (2009) and Rogerson-Revell (2011), who argue that English codas carry not only phonetic information but also important grammatical meanings such as plurality, third-person singular agreement, and past tense. Consequently, inaccurate realization of these final consonants extends beyond segmental pronunciation problems and may affect the comprehensibility during spoken interaction.

The comparatively high error rate for the past tense marker /d/ (58.3%) further demonstrates the interaction between phonological complexity and grammatical processing. Unlike lexical consonants, English regular past tense endings require learners to simultaneously retrieve morphological information and apply context-dependent allophonic rules (/t/, /d/, or /ɪd/). Such dual processing increases cognitive demands during speech production, particularly in spontaneous communication. According to Disney and Le (2024), Vietnamese learners frequently omit or simplify word-final /s/ and /z/ because these sounds contribute little to lexical identification in Vietnamese but play a substantial grammatical role in English. A similar explanation may account for the omission of /d/ observed in the present study. During speech production, learners appear to prioritize lexical meaning over grammatical accuracy, resulting in the deletion or reduction of less perceptually salient morphemes. This observation also aligns with Bui et al. (2021), who reported that Vietnamese university students consistently omitted or substituted English final consonants despite several years of formal English instruction.

Table 4

Errors in pronouncing final consonant sounds (reading passage (RP) and free speech (FS))

Final sounds	Errors
/**z**/ Speaker 1: **clings** (RP, 0:19), **quivers** (RP, 0:31), **plays** (RP, 0:35), **couples** (FS, 0:45) Speaker 2: **colors** (RP, 0:12), **gods** (RP, 0:59)	/**s**/ was produced instead
/**ɪz**/ Speaker 2: **increases** (RP, 1:43), **causes** (RP, 1:29), **enhances** (FS, 0:16)	/**s**/ was produced instead
/**t**/ Speaker 1: **utmost**, **respect**, **pronounced** (RP, 0:27, 0:28, 0:24) Speaker 2: **sunlight** (RP, 0:02), **caused** (RP, 1:17), **passed** (RP, 1:08), **paste** (FS, 0:09)	/**t**/ was unaspirated
/**d**/ Speaker 1: **urged** (RP, 0:52) Speaker 2: **formed** (RP, 1:34), **colored** (RP, 1:42), **observed** (RP, 1:47)	/**d**/ was omitted

With regard to L1, there are no equivalent sounds for /tʃ/ in Vietnamese. As Nation and Newton (2020) proposed that the sound system of learners' first language literally influence their second language sounds. Additionally, L2 learners can find difficulty in articulating L2 phonemes which are not in L1. Therefore, it is understandable that the two participants cannot pronounce /tʃ/ correctly. The same situation happens for producing final sounds of s rules and -ed rules as Vietnamese does not have ending sounds system as English. Therefore, Vietnamese learners struggle to pronounce these final codas accurately.

4.2. Word stress and Intonation

Depending on the recordings, it is obvious that learners mostly did not make stress on words as well as give an inaccurate English speech melody. It can be seen that learners usually omitted word stress, made wrong tone groups division and incorrect tonic placement during speech.

**Examples for incorrect word stress: (Similar patterns happen for L2 Speaker 2)*

Native speakers	L2 Speaker 1
During reading passage (RP): Ancient /ˈeɪnʃənt/ (0:10) Several /ˈsevərəl/ (0:13) Feeling /ˈfiːlɪŋ/ (0:24)	She made no stress without any pitch change. /eɪnʃənt/ /sevərəl/ /filɪŋ/
During free speech (FS):	She made no amplitude for the stress parts.

Music /ˈmjuːzɪk/ (0:12)	/mjuːzɪk/
Melody /ˈmelədi/ (0:32)	/meloʊdi/

**Examples for incorrect speech melody (L2 Speaker 1):*

The tonic (phrasal stress) marked by pitch movement:

The thought groups marked by: //

During reading passage: (RP Speaker 1, 0:17)

Native speaker: A **long**,// **flowing beard clings** to his **chin**,// giving those// who **observe** him a **pronounced feeling** of the **utmost respect**.//

L2 Speaker 1: A long, flowing **beard clings** to his **chin**,// giving those// who observe him a pronounced feeling of the// **utmost**// **respect**.

During free speech: (FS Speaker 1, 1:03)

Native speaker: It **shows** that// he **wants** everyone to **have respect** for everybody//because **everybody** is **equal**.//

L2 Speaker 1: It shows that he wants// everyone to have// uhm **respect** for everybody// because everybody is equal.//

As the analysis above, the participants usually make incorrect tonic stress and the phrasal stress. It can be seen that the speaker gave the incorrect tone groups division, although there is a comma between 'a long' and 'flowing beard', she did not make the pause for it. Furthermore, it seems that the learner used the same tone to speak for both reading a loud and free speech, and she did not really make any rise-fall movement. In this case, the sounds the speaker produced are likely to be affected by the L1 sounds' system as she spoke without any intonation. According to Nation and Newton (2020), age is one of the factors affecting the learning of another sound system. Besides, when learners get older, their first language is more stable, which causes the disruption in their perception of second language. Therefore, these learners have the tendency to put their Vietnamese tones in speaking English words or producing English prosody since they are adults and advanced in L1. As a result, learners did not put stress placement into words and produce incorrect intonation. This will cause the misunderstanding during conversation as well as the providing of an incomprehensible speech.

4.3. Learners' pronunciation goals and settings

Both participants reported a desire to achieve native-like pronunciation for future employment opportunities and international communication. Although these goals reflect learners' personal aspirations, contemporary pronunciation research suggests that intelligibility and comprehensibility may represent more realistic instructional objectives than native-like pronunciation (Derwing & Munro, 2005). Consequently, pronunciation instruction should prioritize features that contribute most directly to successful communication, particularly final

consonants, word stress, and intonation. As the analysis below, it seems that learners can communicate with English native speakers to some extent. Although the learners have studied English for more than 10 years, they still have troubles in their pronunciation and moderate level of listening comprehension. As the description above, the listeners may find some difficulties to be fully understood the messages they conveyed. Therefore, the learners still need to practice more as well as find some appropriate methods for further development in their pronunciation to achieve their goals.

4.4. Limitations

Several limitations are acknowledged during the analysis in this study. First, there were only two participants in this research, which limits the generalizability of the broader population of Vietnamese EFL learners. Despite allowing detailed phonological analysis from the speech recordings, this study acknowledge the lack of other additional data sources, such as interviews, classroom observations or listener judgements. Therefore, further studies should implement these methods to gain deeper insights into the causes and consequences of challenges that second language learners encounter when articulating these pronunciation features. Additional limitations include the absence of inter-rater reliability procedures, the lack of acoustic analysis, and the absence of listener judgments regarding comprehensibility and other aspects. As pronunciation analysis may involve subjective interpretation, future studies should incorporate multiple raters and listener-based measures to strengthen analytical rigor.

Second, this study focused on some specific pronunciation aspects, namely consonants, words stress and intonation. There are other pronunciation features that need considering such as vowels, rhythm, speech rate, and connected speech. Hence, further research should take these aspects into account to extend the scope of current investigation.

Finally, the study adopted a descriptive approach and did not examine the effectiveness of any pronunciation intervention. Future studies may implement instructional interventions as well as pedagogical practices mentioned below to determine whether targeted pronunciation training can improve Vietnamese learners' pronunciation of final consonants and prosodic features.

5. Pedagogical implications

With the analysis and discussion above, several pedagogical implications emerge. As these learners had difficulties pronouncing final sounds, practical tasks with explicit instructions on final consonants should be carefully designed and implemented during speaking courses. According to (Knight, 2012; Nguyen & Newton, 2020), greater emphasis should be made to final sounds to support communicative effectiveness. Therefore, teachers should provide explicit instruction regarding the pronunciation rules of /s/, /z/, /ɪz/, /t/, /d/, and /ɪd/, followed by extensive communicative practice. Activities such as minimal-pair discrimination, information-gap tasks, dictation exercises, and peer correction can help learners develop greater awareness of these sounds. Some of these activities were included in the feedback (see Appendix C) for these learners to improve their pronunciation after taking part in this study.

Regarding practices for word stress and intonation improvements, previous researchers have suggested that suprasegmental instruction contributed significantly to these pronunciation features (Derwing & Munro, 2005; Yates, 2017). As described in the findings, both participants struggled to stress English words correctly and have appropriate intonation, leading to the need

of integrating stress placement, tonic stress, and pitch movement tasks into their current speaking course (see Appendix C).

As for further development in this field, pronunciation instruction should be integrated into communicative speaking activities rather than taught in isolation. By embedding pronunciation practices within meaningful interactions, learners can develop greater control over pronunciation features while simultaneously improving their communicative effectiveness.

6. Conclusion

In summary, this qualitative case study investigated the pronunciation features of two Vietnamese EFL sophomores through a read-aloud task and a free-speaking task. Drawing on the phonological frameworks of Roach (2009), Rogerson-Revell (2011), and Cruttenden (2014), the study examined consonant production, word stress, and intonation patterns. The findings revealed that the participants had the tendency of omitting the final sounds and had troubles in producing a good English intonation. Despite articulating many English consonants and consonant clusters correctly to some extent, the learners experienced persistent difficulties with final consonants, particularly those associated with plural, third-person singular, and past tense verbs. In addition to this, both learners demonstrated limited control of English word stress and intonation, frequently producing speech with reduced pitch variation and inaccurate stress placement. These errors have made an impact on the communicative effectiveness and raise the awareness of the importance of final sounds articulation. These pronunciation patterns may have implications for communicative effectiveness; however, further research incorporating listener judgments is needed to verify this relationship. The findings of this study also align with previous studies (Disney & Le, 2024; Bui et al., 2021; Minh et al., 2024), calling for further research in this field. To support Vietnamese sophomores in improving their pronunciation, this paper provided some pedagogical implications to raise awareness of L2 Vietnamese learners about these two aspects of pronunciation. Therefore, it is necessary for both learners and teachers to pay attention to these pronunciation features and have appropriate methods for improvements.

References

Al-khresheh, M. H. (2024). Phonetic challenges in English: the impact of mispronunciation of the bilabial plosive/p/on communication among Saudi EFL learners. *Cogent Arts & Humanities*, *11*(1). https://doi.org/10.1080/23311983.2024.2390777

Bui, T. T. L., Mai, T. H., & Diep, H. N. (2021). COMMON ERRORS IN PRONOUNCING FINAL CONSONANTS OF ENGLISH-MAJORED SOPHOMORES AT TAY DO

UNIVERSITY, VIETNAM. *European Journal of English Language Teaching*, *6*(3). https://doi.org/10.46827/ejel.v6i3.3640

Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). New York: Routledge.

Cummins, F. (2015). Rhythm and Speech. In *The Handbook of Speech Production*, M.A. Redford (Ed.). https://doi.org/10.1002/9781118584156.ch8

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders* (3rd ed.). Philadelphia, PA: W.B. Saunders Company.

Derwing, T. & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*(3), 379-398.

Disney, S. J., & Nu Cam Le, L. (2024). Word-Final /s/-/z/ Omission in Vietnamese English. *Languages*, *9*(10), 327. https://doi.org/10.3390/languages9100327

Fairbanks, G. (1960). *Voice and articulation drillbook, 2nd edn*. New York: Harper & Row. pp124-139.

Hien, T. T. T. (2024). Common Mistakes in Pronunciation among Students in The English for Specific Purposes Department, Hanoi University. *Vietnam Journal of Social Sciences & Humanities*, *10*(5), 488. https://doi.org/10.33100/tckhxhnv10.5.TaThiThanhHien

Knight, Rachael-Anne (2012). *Phonetics: A coursebook*. New York: Cambridge University Press.

Minh, B. N. N., Lian, A., & Vy, L. T. M. (2024). Improving the Pronunciation of Voiceless Consonant Sounds in Vietnamese EFL Undergraduates Using the Simplified Verbotonal Approach. *Theory and Practice in Language Studies*, *14*(9), 2690–2701. https://doi.org/10.17507/tpls.1409.05

Nation, I.P.S. & Newton, J. (2020). *Teaching speaking and listening* (2nd ed.). New York: Routledge.

Nguyen, L. T., & Newton, J. (2020). Pronunciation Teaching in Tertiary EFL Classes: Vietnamese Teachers' Beliefs and Practices. *The Electronic Journal for English as a Second Language*, *24*(1),1-20.

Phạm, B., & McLeod, S. (2016). Consonants, vowels and tones across Vietnamese dialects. *International Journal of Speech-Language Pathology*, *18*(2), 122–134. https://doi.org/10.3109/17549507.2015.1101162

Pourhosein Gilakjani, A. (2016). What Factors Influence the English Pronunciation of EFL Learners? *Modern Journal of Language Teaching Methods (MJLTM)*, *6*(2), 314-326.

Roach, P. (2009). *English Phonetics and Phonology* (4th ed.). Cambridge: Cambridge University Press.

Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. New York: Continuum International Publishing Group.

Yates, L. (2017). Learning how to speak: Pronunciation, pragmatics and practicalities in the classroom and beyond. *Language Teaching, 50*(2), 227-246.

Appendix A

English

Vietnamese equivalent sounds

Consonants	English	Vietnamese
/**k**/	**C**at	**C**át
/**m**/	**M**om	**M**ẹ
/**n**/	**N**o	**N**ói
/**s**/	**S**ound	**X**e

/**r**/	**R**oad	**R**ẽ
/**h**/	**H**at	**H**át
/**g**/	**G**ood	**G**ót
/**f**/	**F**at	**Ph**ở
/**b**/	**B**oots	**B**ò
/**w**/	**Wh**at	**Q**uả

Appendix B

Free speaking task: Questions

1. Tell me what is your favorite song
2. Can you tell me more about it?
3. Have you eaten any weird food?
4. How was your feeling when you ate these kinds of food?

Read-aloud task:

The Rainbow Passage

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is , according to legend, a boiling pot of gold at one end.

People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

Throughout the centuries people have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Others have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbows. Many complicated ideas about the rainbow have been formed.

The difference in the rainbow depends considerably upon the size of the drops, and the width of the colored band increases as the size of the drops increases. The actual primary rainbow observed is said to be the effect of super-imposition of a number of bows. If the red of the second bow falls upon the green of the first, the result is to give a bow with an abnormally wide yellow band, since red and green light when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green or blue.

The Grandfather Passage

You wished to know all about my grandfather. Well, he is nearly ninety-three years old; he dresses himself in an ancient black frock coat, usually minus several buttons; yet he still thinks as swiftly as ever. A long, flowing beard clings to his chin, giving those who observe him a pronounced feeling of the utmost respect. When he speaks, his voice is just a bit cracked and quivers a trifle. Twice each day he plays skilfully and with zest upon our small organ. Except in the winter when the ooze or snow or ice prevents, he slowly takes a short walk in the open air each day. We have often urged him to walk more and smoke less, but he always answers, “Banana oil!” Grandfather likes to be modern in his language.

Appendix C

Feedback and recommendation for learners

Overall, your pronunciation is clear but it still has some problems that need to be addressed. Firstly, let’s look at your pronunciation strengths:

- Almost of the consonants are pronounced correctly, both voiced or voiceless.
- 70-80% of your vowels are articulated well
- The consonants clusters are produced well

As mentioned below, it is good for you to keep up these strong points in your pronunciation and practice more to develop your voice like native speakers. However, there are weaknesses about your pronunciation that you also need to work on:

- Incorrect final sounds for s rules pronunciation (/s/, /z/, /iz/) and past-tense regular verbs pronunciation (/t/, /d/, /id/)
- Incorrect word stress
- Incorrect intonation

It is good for you to listen to your recordings by yourself again and notice the errors mentioned above. You can see that you have produced only /s/ as final sounds for all the words ending with s. Additionally, you have omitted the final sounds for past tense verbs and some word stress. Besides, there is no pitch change of English during your speech and you are likely to use same tone when producing speech. Therefore, this may cause misunderstanding when you communicate with English native speakers. Hence, it is necessary for you to carry out these problems for your further pronunciation development.

Here are some activities that are suggested to help you improve your pronunciation:

1. Final sounds:

Activity: *Work with your friends. Listen to your partner speaking a group of words and find out the word with different final sounds.*

plays hands words peas buses hills

cats sleeps sits hats picks drops

You and your friends take turns, create a piece of paper with groups of words like this beforehand and keep practice together.

2. Word Stress:

Activity:

a. Listen to the following words and write the number of syllables in each word

Considerably ____ Social ____

Explanation ____ Understandable ____

Breakable ____ Expectancy ____

Immediately ____ Melody ____

b. Listen again and circle the stressed syllable in each word. Then practice saying these words to your partner. Check your partner's stress.

This activity will help you to get used to making word stress. You and your friends should take turns and practice together.

3. Intonation:

Activity: *You and your friend have been shopping. Make a conversation about what you purchased using the information given below:*

A: What / you / got?
B: I / buy / shoes / two pairs / jeans / hat / T-shirt / this bag.
A: Quite a lot. I only / buy / couple / T-shirt / sneakers / my upcoming vacation.
B: You / buy / that jacket?
A: Which one? One / I / try on / Tommy's?

Pay attention about your voice when to rise and fall during practicing. After that, you and your partner should check your intonation through watching a video which native speakers do the same scenario. This activity literally can help you have a better intonation.

#136 THE PEDAGOGICAL IMPACT OF ARTIFICIAL INTELLIGENCE TOOLS ON ENGLISH MAJORS' WRITING SKILLS AT HO CHI MINH CITY UNIVERSITY OF FOREIGN LANGUAGES - INFORMATION TECHNOLOGY

Nguyễn Minh Tuấn [1]

Abstract

The increasing use of Artificial Intelligence (AI) tools in education has begun to reshape writing instruction in English language learning contexts. However, there is still limited research on how English-major students in Vietnam perceive the use of these tools, particularly at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). This study explores third-year English majors' perceptions and self-reported experiences regarding the use of AI tools in writing through a mixed-methods approach. Data were collected from a questionnaire completed by 169 students and semi-structured interviews with eight participants. The findings indicate that many students perceive AI tools as beneficial for supporting various aspects of writing, particularly grammar, vocabulary, and idea organization. Participants also reported that AI tools helped them generate ideas, identify errors, and revise their drafts more efficiently. At the same time, some students expressed concerns that excessive reliance on AI tools could reduce opportunities for independent thinking and negatively affect their confidence in writing. Overall, the study suggests that students generally view AI tools as valuable support for the writing process when used appropriately. The findings highlight the

[1] *M.A., Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology, email: tuannm@huflit.edu.vn*

importance of promoting a balanced and guided integration of AI tools in English writing instruction.

Keywords: Artificial Intelligence tools, writing skills, English majors, language learning

1. Introduction

1.1 Background and Statement of the Problem

In recent years, Artificial Intelligence (AI) has become more common in education, especially in English language learning. Many students now use AI-powered applications such as ChatGPT, Gemini, Grammarly, and QuillBot to support their writing. These tools can help users correct grammar mistakes, improve vocabulary, generate ideas, and receive quick feedback during the writing process.

Writing is often considered one of the most difficult skills in learning English because students need not only good grammar and vocabulary but also the ability to organize ideas clearly and think critically. For third-year English majors, writing is especially important because they are required to complete more advanced academic tasks such as essays, reports, and research assignments. Because of this, many students use AI tools to assist them in different stages of writing, from brainstorming ideas to revising and editing their work.

Recent studies indicate that AI-assisted writing tools can support language learners in several aspects of writing. From students' perspectives, these tools are often considered useful for improving grammar, expanding vocabulary, generating ideas, organizing ideas, and providing immediate feedback during the writing process. However, most existing research tends to focus mainly on the advantages of using AI tools. In comparison, fewer studies have examined students' perceptions of possible challenges and risks, such as overreliance on AI-generated content, reduced opportunities for independent thinking, loss of creativity, and lower confidence in their own writing skills. In addition, although research on AI-assisted writing has been growing internationally, there is still a lack of studies focusing specifically on English-major students in Vietnamese universities.

Although these technologies can make writing more convenient and time-saving, there are still concerns about their influence on students' actual writing ability. Some students believe AI tools help improve the quality of their writing, while others worry that depending too much on technology may reduce independent thinking, creativity, and confidence in writing. In addition, research on the impact of AI tools on English-major students in the Vietnamese university context is still limited, especially at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT).

1.2 Research Objectives and Questions

This study aims to investigate the pedagogical impact of AI writing tools on the writing skills of third-year English majors at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). In particular, the study explores how often students use AI tools in writing tasks, how these tools affect different aspects of writing skills, and whether AI usage influences students' confidence in writing.

To achieve these objectives, the study addresses the following research questions:

1. To what extent do AI tools influence students' writing skills in terms of grammar, vocabulary, coherence, and idea development?
2. How frequently do third-year English majors use AI tools in writing tasks?
3. How do AI tools influence students' confidence and learning experiences in English writing?

1.3 Significance of the study

This study is expected to contribute both theoretically and practically to the field of English language education. Theoretically, it provides more information about how AI tools influence the writing skills of English-major students in the Vietnamese university context.

Practically, the findings may help students better understand both the advantages and possible disadvantages of using AI tools in writing. The study may also provide useful suggestions for teachers on how to integrate AI tools into writing instruction in a balanced and effective way. In addition, the findings may serve as a useful reference for future research related to AI-assisted language learning and writing instruction.

1.4 Scope of the Study

The study examines students' use of AI-assisted writing applications such as ChatGPT, Gemini, Grammarly, QuillBot, and other similar tools. It mainly explores how these technologies influence students' writing performance in terms of grammar, vocabulary, coherence, organization, and idea development.

The participants of the study are limited to third-year English-major students at HUFLIT. Therefore, the findings may not be generalized to students from other majors or educational institutions.

2. Literature Review

2.1 AI tools in English writing

2.1.1 Definition and types of AI writing tools

In recent years, AI tools have become increasingly common in English language learning, especially in writing. AI writing tools are applications that use artificial intelligence to help learners during the writing process. Some popular examples are Grammarly, QuillBot, and ChatGPT. These tools can support students by checking grammar, suggesting vocabulary, paraphrasing sentences, generating ideas, and giving feedback on writing.

Many students use AI tools at different stages of writing. For example, students may use them to brainstorm ideas before writing, improve sentence structures while drafting, or correct grammatical mistakes after finishing a paragraph. Because of these functions, AI writing tools are now widely used by university students to make writing tasks easier and more convenient.

2.1.2 Theoretical perspectives on AI-Assisted writing

The use of AI tools in writing can be explained through several learning theories. One important theory is constructivism. According to Hausfather (1996), learners develop knowledge through interaction, guidance, and feedback. In AI-assisted writing, students interact with AI tools and receive suggestions or corrections that may help them improve their writing skills.

Another related theory is cognitive learning theory. Sweller (1988) explains that reducing cognitive load can help learners focus more on higher-level thinking tasks. In writing,

AI tools may reduce the pressure of lower-level tasks such as grammar correction or spelling, allowing students to focus more on idea development, organization, and coherence.

These theories suggest that AI tools can support the learning process and improve writing performance. However, the effectiveness of AI-assisted writing still depends on how students use these tools in their learning.

The theoretical perspectives discussed above provide a foundation for the variables investigated in this study. According to constructivist theory, AI tools can function as supportive learning resources that provide feedback and guidance, enabling students to participate more actively in the writing process. This perspective is reflected in the study's examination of students' perceptions of writing support, feedback, confidence, and overall learning experiences. In addition, cognitive learning theory suggests that AI tools can help reduce the cognitive load associated with lower-level writing tasks, such as grammar checking and vocabulary selection, allowing students to focus more on developing and organizing their ideas. However, excessive dependence on AI-generated assistance may negatively affect students' autonomy, creativity, and ability to solve writing problems independently. Therefore, the questionnaire and interview questions were developed to explore both the perceived benefits and the potential concerns related to the use of AI-assisted writing tools.

2.1.3 Benefits of AI tools in writing

Many researchers have found that AI tools can positively support students' writing development. Kessler (2018) states that automated feedback systems help learners improve grammatical accuracy because students can quickly identify and correct mistakes in their writing. Similarly, tools such as Grammarly and ChatGPT may help students improve vocabulary use, sentence structure, and overall writing organization.

Several studies also show that AI tools can support students during the revision process. Hwang et al. (2023) and Seufert et al. (2021) found that AI-generated feedback helps learners revise drafts and organize ideas more effectively. Since students can review and edit their work independently, AI tools may also encourage learner autonomy.

Besides improving writing quality, AI tools may also increase students' motivation and engagement in learning. Kukulska-Hulme and Shield (2008) suggest that technology-supported learning environments encourage independent learning outside the classroom. Similarly, Zhai et al. (2024) argue that AI-supported educational tools can increase students' participation and motivation because they provide quick feedback and personalized support.

Other researchers, such as Nazari et al. (2021) and Utami and Winarni (2023), found that students who regularly use AI tools often feel more confident when completing writing tasks. This suggests that AI tools may support not only writing performance but also students' attitudes toward learning English writing.

2.1.4 Challenges of AI tools in writing

Although AI tools provide many advantages, researchers have also identified several challenges related to their use. One major concern is that students may become too dependent on AI tools. Zhai (2022) argues that overreliance on AI may reduce students' critical thinking ability and limit their ability to express original ideas independently.

Similarly, Ranalli (2018) suggests that when learners rely too much on AI support, they may become less able to evaluate their own writing without technology. Zou et al. (2023) and

Yan (2023) also warn that excessive use of AI tools may weaken independent writing ability and increase the risk of plagiarism.

Another issue is that AI-generated feedback may not always be fully accurate or detailed. According to Su et al. (2023), AI feedback often focuses mainly on grammar and vocabulary rather than deeper aspects of writing such as creativity, idea development, or contextual meaning. Therefore, AI tools cannot completely replace teacher feedback in the writing classroom.

In addition, Warschauer and Grimes (2008) point out that too much dependence on automated systems may reduce interaction between teachers and students, which plays an important role in developing communication and critical thinking skills. Syuhra, Chandra, and Rosalina (2025) further explain that students need proper guidance when using AI tools because misuse of these technologies may negatively affect learning outcomes.

2.2 Previous studies on AI tools and writing skills

Many previous studies have explored the relationship between AI tools and English writing skills. Most studies show that AI-assisted writing can help students improve grammar, vocabulary, and organization in writing. Researchers generally agree that immediate feedback from AI tools helps students revise their writing more efficiently and produce more accurate texts.

At the same time, some studies have discussed the possible negative effects of AI-assisted writing. Several researchers argue that students who rely too heavily on AI tools may become less independent in writing and may not fully develop critical thinking skills. In addition, some studies suggest that AI-generated feedback still cannot fully replace teacher feedback because it may lack contextual understanding and detailed explanation.

Previous research has also shown that AI tools may positively affect students' motivation, confidence, and engagement in learning. Many students consider AI tools useful because they provide quick support and personalized feedback during writing tasks. However, most previous studies focus only on specific aspects of writing, such as grammar or vocabulary, and only a limited number of studies combine writing performance with students' personal experiences and opinions.

Therefore, more research is still needed to better understand how AI tools influence students' writing skills in different educational contexts, especially among English-major students in Vietnam.

2.3 Research Gap

Although previous studies have highlighted various benefits of AI-assisted writing, such as support for grammar improvement, vocabulary development, feedback, and revision, relatively few studies have examined both the advantages and potential drawbacks of AI use within the same research framework. In particular, limited research has explored students' perceptions of how AI tools affect their confidence, autonomy, creativity, and reliance on technology during the writing process.

Moreover, studies focusing on English-major students in Vietnamese universities are still limited, especially those using a mixed-method approach to investigate both quantitative trends and students' personal experiences. As AI-assisted writing tools are becoming increasingly common in higher education, it is important to gain a deeper understanding of

students' perceptions to support the effective and responsible integration of AI into language learning.

Therefore, this study aims to investigate English-major students' experiences and perceptions of AI-assisted writing by examining four related dimensions: AI tool use, perceived writing support, perceived concerns, and students' confidence and learning experiences.

3. Research Methodology

3.1 Research design

This study used a mixed-method research design to investigate the impact of AI tools on the writing skills of third-year English-major students at HUFLIT. The study also explored students' perceptions and experiences of using AI tools in writing activities.

A mixed-method approach was considered appropriate because the study aimed not only to examine students' writing performance through numerical data but also to understand their attitudes, opinions, and experiences related to AI-assisted writing. By combining quantitative and qualitative methods, the researcher was able to obtain a broader understanding of how AI tools influence students' writing skills.

3.2 Participants

The questionnaire survey involved 169 third-year English-major students at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT). These students had already completed several writing courses and were familiar with using AI tools in academic writing tasks. Convenience sampling was used in this study because the participants were easily accessible during the data collection period and voluntarily agreed to take part in the research. While this method allowed for efficient data collection, it may also lead to selection bias, as those who chose to participate might differ from those who did not. Therefore, the findings should be interpreted carefully and cannot be generalized to all English-major students in Vietnam. Instead, the results mainly reflect the perceptions and experiences of third-year English-major students at HUFLIT.

For the qualitative phase, eight students were selected to participate in semi-structured interviews. The interviews were conducted to gain a deeper understanding of students' experiences and perceptions of AI-assisted writing. The participants shared their opinions about the advantages and challenges of using AI tools in their writing process.

3.3 Research instruments

This study used two main research instruments: questionnaires and semi-structured interviews. The combination of these instruments allowed the researcher to collect both quantitative and qualitative data.

According to Wasti et al. (2022), mixed-method research helps researchers gain a more comprehensive understanding of a research problem by combining statistical data with personal experiences and opinions. Therefore, questionnaires were used to collect general information and numerical data, while interviews were used to gather more detailed insights from participants.

3.3.1 Questionnaire

The questionnaire was used to collect information about students' experiences and perceptions of using AI tools in writing. Questionnaires are commonly used in educational research because they allow researchers to collect data from a large number of participants

within a short period of time (Breakwell, Barnett, & Wright, 2020). The questions mainly focused on students' frequency of AI tool usage, their attitudes toward AI-assisted writing, and their perceptions of how AI tools affect their writing skills.

A pilot study involving 10 students was conducted prior to the main survey. Based on the participants' feedback, several items in the questionnaire were revised to improve clarity, simplify the wording, and reduce potential ambiguity. These adjustments helped ensure that respondents could better understand the questions and provide more accurate responses. Although questionnaires are useful for collecting statistical information, they may sometimes limit the depth of responses because participants often provide short answers (Kumar, 2023). Therefore, interviews were included to support and expand the quantitative findings.

3.3.2 Semi-Structured interviews

After the questionnaire survey, semi-structured interviews were conducted with eight participants. The purpose of the interviews was to explore students' experiences and opinions in greater detail. According to Creswell (2014), interviews are useful for collecting detailed information through direct interaction with participants. In this study, the interviews helped clarify students' views about the benefits and challenges of using AI tools in English writing.

However, interviews may sometimes be influenced by researcher bias because participants' responses can be affected by the interviewer's reactions or behavior (Kvale, 2009). For this reason, the researcher attempted to remain neutral throughout the interview process.

3.4 Data collection procedures

In the first phase, the questionnaire survey was distributed online through Google Forms over a period of two months. Clear instructions were provided to help participants complete the questionnaire accurately. Participants were also informed that all responses would remain confidential.

In the second phase, follow-up interviews were conducted with eight students who had completed the questionnaire survey. They were conducted in person at the university to encourage direct interaction between the researcher and participants. With participants' permission, the interviews were audio-recorded and video-recorded for research purposes. The interview data were later transcribed and analyzed to support and explain the quantitative findings more clearly.

3.5 Data analysis procedures

3.5.1 Quantitative data analysis

The quantitative data collected from the questionnaires were analyzed using SPSS Statistics 20. Descriptive statistics such as frequencies, means, and standard deviations were used to summarize participants' responses and identify general trends in the data.

For the Likert-scale questions, responses were coded and analyzed using statistical measures including percentage, mean, and standard deviation. These statistical results helped identify students' perceptions of the usefulness of AI tools in improving their English writing skills.

Cronbach's Alpha analysis was conducted in SPSS to examine the reliability of the questionnaire items. The questionnaire included three main constructs: students' perceptions

of AI in writing, writing development, and concerns. Each construct consisted of five items measured using a five-point Likert scale.

Table 1

Cronbach's Alpha of perceptions' items

Reliability Statistics	
Cronbach's Alpha	Number of Items
0.810	5

Table 2

Cronbach's Alpha of writing developments' items

Reliability Statistics	
Cronbach's Alpha	Number of Items
0.849	5

Table 3

Cronbach's Alpha of concerns' items

Reliability Statistics	
Cronbach's Alpha	Number of Items
0.844	5

As shown in Table 1, the Cronbach's Alpha value for the "perceptions" items was 0.810, indicating that the items had good reliability. In Table 2, the "writing development" items achieved a Cronbach's Alpha value of 0.849, which showed a high level of internal consistency among the items. Similarly, Table 3 presents the Cronbach's Alpha value of the "concerns" items, which was 0.844, indicating good reliability as well.

According to Nunnally (1978), a Cronbach's Alpha value above 0.70 is considered acceptable for research studies. Therefore, all constructs in this study were considered reliable, and all questionnaire items were retained for further data analysis.

3.5.2 Qualitative data analysis

The qualitative data collected from the interviews were analyzed using thematic analysis. First, the researcher reviewed the interview recordings and transcripts carefully. After that, the responses were coded to identify recurring ideas and common themes related to students' experiences of using AI tools in writing.

The themes identified from the interviews were then used to support and explain the quantitative findings. This helped the researcher better understand how AI tools influence the writing skills and learning experiences of third-year English-major students.

4. Findings

4.1 Quantitative findings

Figure 1

The frequency of using AI tools in writing skill

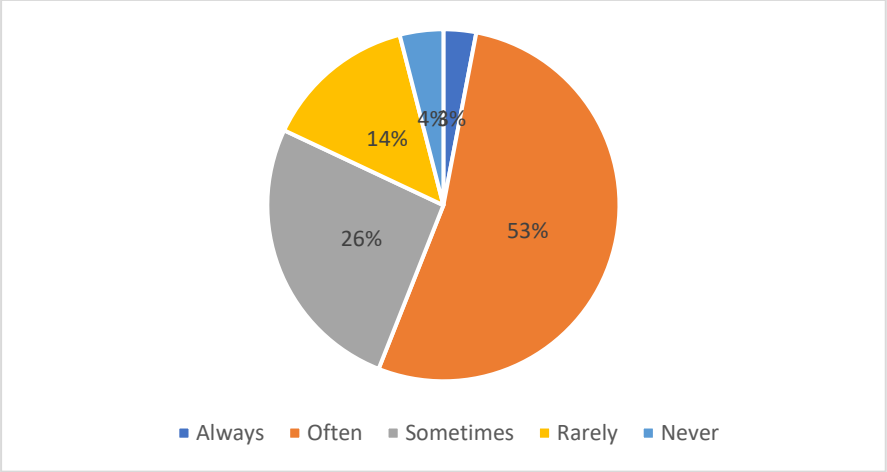


Figure 1 presents the frequency of AI tool usage among students in learning and completing writing tasks. As shown in the findings, most participants reported that they often use AI tools, representing 53% of the responses. In addition, 26% of the respondents indicated that they sometimes use AI tools, while 14% reported rarely using them. Only a small proportion of students were found at the two opposite ends of usage frequency. Specifically, 3% of the participants stated that they always use AI tools, whereas 4% answered that they never use them. Overall, the results show that AI tools are widely adopted among third-year English-major students.

Figure 2

The types of AI tools used in writing skill

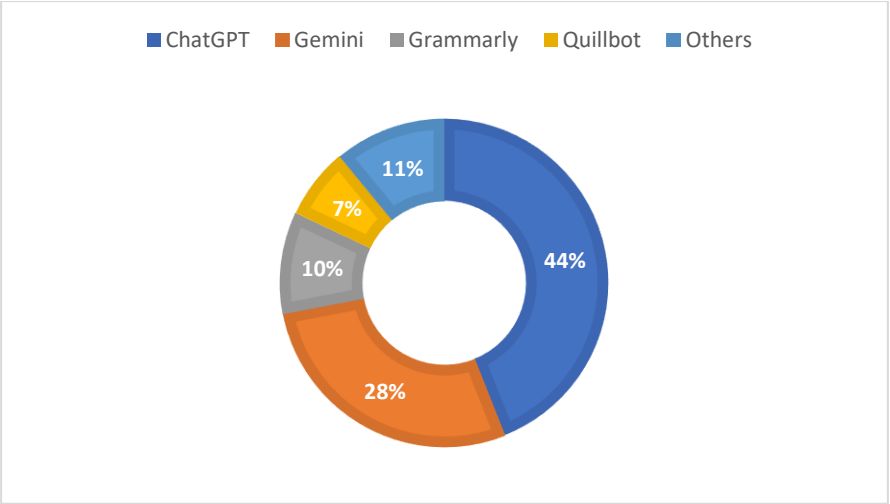


Figure 2 shows the AI tools most frequently used by students in their writing activities. According to the findings, ChatGPT was the most commonly used tool, accounting for 44% of the responses. The second most frequently used tool was Gemini, which represented 28% of the responses. Meanwhile, 11% of the participants reported using other AI tools for writing support. In addition, Grammarly accounted for 10% of the responses, while QuillBot was the least frequently used tool, with only 7%. Overall, the findings indicate that AI-powered writing assistants, especially ChatGPT and Gemini, are widely used among third-year English-major students

Table 4

Students' perceptions of AI tools in writing skill

Students' Perceptions of AI Tools	Degree of agreement					Mean	Stand. Dev.
	SD	D	N	A	SA		
Generating ideas	0%	3%	20.1%	51.5%	25.4%	3.99	0.76
Providing clear, useful feedback	0%	3.6%	18.3%	56.2%	21.9%	3.96	0.74
Learning from personal mistakes	0%	1.8%	20.7%	53.3%	24.3%	4.00	0.72
Saving time	0%	2.4%	26%	50.3%	21.3%	3.91	0.75
Challenging students	0%	1.8%	27.2%	46.7%	24.3%	3.93	0.77

Table 4 presents students' perceptions of AI tools in English writing activities. Overall, the findings show that students had positive perceptions toward the use of AI in writing, with all items receiving relatively high mean scores ranging from 3.91 to 4.00.

Among the five statements, "Learning from personal mistakes" obtained the highest mean score (M = 4.00, SD = 0.72), indicating that students found AI tools useful in identifying and correcting writing errors. In addition, students also showed positive perceptions of AI tools in generating ideas (M = 3.99, SD = 0.76) and providing useful feedback (M = 3.96, SD = 0.74). The findings further revealed that AI tools were considered helpful in saving time and supporting the writing process.

In general, the results suggest that third-year English-major students viewed AI tools as beneficial in supporting various aspects of English writing, especially idea generation, feedback, and error correction.

Table 5

Perceived writing support from AI tools

Impacts of AI Tools	Degree of agreement					Mean	Stand. Dev.
	SD	D	N	A	SA		
Grammar accuracy	0%	1.8%	16.6%	52.7%	29%	4.09	0.72
Vocabulary expansion	0%	2.4%	8.3%	47.9%	41.4%	4.28	0.72
Coherence and organization	0%	3%	14.2%	50.3%	32.5%	4.12	0.76
Idea development	0%	3%	13.6%	56.8%	26.6%	4.07	0.72
Writing process	0%	4.7%	14.2%	53.8%	27.2%	4.04	0.79

Table 5 indicates students' experience of the impact of AI tools on the development of English writing skills. Overall, the findings indicate that students had positive perceptions toward the use of AI tools in writing, as all aspects received relatively high mean scores ranging from 4.04 to 4.28.

Among the five aspects, vocabulary expansion achieved the highest mean score (M = 4.28, SD = 0.72), suggesting that students perceived AI tools as helpful for supporting vocabulary use during the writing process. In addition, students also reported positive perceptions regarding coherence and organization (M = 4.12, SD = 0.76), grammar accuracy (M = 4.09, SD = 0.72), and idea development (M = 4.07, SD = 0.72). The writing process received the lowest mean score among the five aspects (M = 4.04, SD = 0.79), although the result still reflected a generally positive attitude.

In general, the findings suggest that third-year English-major students generally perceived AI tools as valuable sources of support across different stages of the writing process, especially vocabulary use, grammar improvement, and writing organization.

Table 6

Students' concerns regarding AI-Assisted writing

Concerns of using AI Tools	Degree of agreement					Mean	Stand. Dev.
	SD	D	N	A	SA		
Dependence on AI tools	0%	3%	13.6%	49.1%	34.3%	4.15	0.76
Reduced writing practice	0%	1.8%	14.8%	54.4%	29%	4.11	0.71
Increased writing anxiety	0%	1.2%	15.4%	46.7%	36.7%	4.19	0.73
Reduced learning autonomy	0%	3%	12.4%	47.3%	37.3%	4.19	0.76
Reduced creativity	0%	3.6%	8.9%	53.8%	33.7%	4.18	0.74

Table 6 shows students' concerns about the use of AI tools in English writing activities. Overall, the findings indicate that students were aware of several possible negative effects of AI-assisted writing, as all items received relatively high mean scores ranging from 4.11 to 4.19.

Among the five concerns, "Increase mistake anxiety" and "Influence learning autonomy" achieved the highest mean scores (M = 4.19), suggesting that many students were worried about becoming less confident and more dependent on AI support during the writing process. Furthermore, students also expressed concerns about weakened creativity (M = 4.18, SD = 0.74) and dependence on AI tools (M = 4.15, SD = 0.76). Although "Decrease regular practice" received the lowest mean score (M = 4.11, SD = 0.71), the result still reflected a relatively high level of concern among the participants.

In general, the findings suggest that while third-year English-major students recognized the benefits of AI tools in writing, they were also concerned about several potential disadvantages, particularly dependence on technology, reduced learning autonomy, and weakened creativity.

4.2. Qualitative findings

The qualitative findings are presented to further explain and support the quantitative results discussed in Section 4.1. The interview data provide a deeper understanding of students' perceptions of AI-assisted writing by exploring both the benefits and concerns associated with

the use of AI tools. By offering detailed explanations and personal experiences, the qualitative findings complement the survey results and contribute to a more comprehensive interpretation of the study's findings.

Theme 1: *AI tools help improve students' writing quality*

The interview findings showed that AI tools play an important role in helping students improve the quality of their writing. Many participants mentioned that AI tools such as ChatGPT, Gemini and Grammarly help them identify grammar mistakes, improve vocabulary use, and make their sentences clearer and more natural. Students also explained that AI-generated feedback helps them revise their writing more effectively and learn from their mistakes.

One participant stated "*AI tools correct grammar mistakes and suggest more academic words and natural expressions. Because of that, my writing becomes clearer and more structured*." (Participant 1)

Similarly, another student told "*They provide suggestions for clearer expressions and more formal vocabulary. Because of that, I can gradually improve my writing ability*." (Participant 2)

Another participant also mentioned "*AI tools help me improve vocabulary and sentence structure, and I can learn from my mistakes*." (Participant 6)

These responses explain the quantitative results presented in Table 5, in which grammar accuracy and vocabulary expansion received relatively high mean scores. While the survey results show that students generally viewed AI tools as helpful for writing-related tasks, the interview data provide specific examples of how students used AI-generated feedback to identify language errors, revise their drafts, and improve their writing. As a result, the qualitative findings offer a deeper understanding of the factors contributing to students' positive perceptions of AI-assisted writing support.

Theme 2: *AI tools support brainstorming and idea development*

Another important finding from the interviews was that AI tools help students generate and develop ideas during the writing process. Many participants explained that they often use AI tools when they have difficulty starting a writing task or when they need additional information and ideas related to a topic.

According to the participants, AI tools can provide useful suggestions, examples, and different perspectives, which help them understand the topic better and organize their ideas more clearly before writing.

One participant explained "*I usually use ChatGPT to generate ideas when I don't know how to start my writing.*" (Participant 1)

Another student stated "*I use AI tools when I need help with ideas or when I don't understand a writing topic clearly*." (Participant 3)

Similarly, another participant shared "*AI tools help me have more information for my ideas, and sometimes their ideas are better than mine*." (Participant 6)

These responses suggest that AI tools are especially helpful during the early stages of writing. They help students overcome difficulties in brainstorming, develop ideas more effectively, and feel more confident when starting their writing tasks.

Theme 3: *Concerns about overreliance on AI tools*

Besides the positive effects of AI-assisted writing, several participants also expressed concerns about becoming too dependent on AI tools. Some students believed that frequent use of AI could reduce their independent thinking, creativity, and confidence when writing without technological support.

One participant mentioned "*Sometimes I depend too much on AI tools, especially when I do not have ideas for writing.*" (Participant 4)

Another student shared "*When I use AI too often, I feel less confident writing by myself because I always want to check AI suggestions first*." (Participant 7)

Similarly, another participant explained "*AI tools are helpful, but sometimes I feel that my creativity becomes weaker because I rely on their ideas too much*." (Participant 5)

These findings indicate that although students recognized many benefits of AI tools, they were also aware of several possible negative effects related to overreliance on technology. Some participants felt that excessive use of AI tools could reduce regular writing practice and weaken their ability to think and write independently. The findings also suggest that students still value creativity and independent learning in the writing process, even when using AI-assisted technologies.

5. Discussion

5.1 Positive impacts of AI tools on writing skills

The findings of this study indicate that most third-year English-major students had positive perceptions toward the use of AI tools in writing activities. Many students believed that AI tools helped improve important aspects of writing, especially grammar accuracy, vocabulary use, coherence, and idea development. Among these aspects, vocabulary improvement received the highest mean score, which suggests that students found AI tools particularly useful for choosing more appropriate and academic vocabulary in their writing.

The finding that vocabulary expansion received the highest mean score may be explained by the direct and noticeable support that AI tools provide during writing. Compared with more complex writing skills, such as idea development and organization, vocabulary improvement is often easier for students to recognize and evaluate. AI tools can quickly suggest synonyms, academic vocabulary, and alternative expressions, enabling students to make immediate changes to their writing. Therefore, students may perceive vocabulary-related support more positively than other types of writing assistance.

Students' positive perceptions of AI-generated feedback may also be related to its accessibility and immediate availability. Unlike traditional feedback from teachers, which may take time to receive, AI tools can provide suggestions instantly while students are writing. This immediate feedback can help students identify language errors and revise their work more efficiently. As a result, students may view AI feedback as a useful and convenient source of support, especially when working on academic writing tasks independently.

The interview findings also supported the quantitative results. Several participants mentioned that AI tools such as ChatGPT, Grammarly, and Gemini helped them correct grammatical mistakes, improve sentence structure, and organize ideas more clearly. In addition, many students explained that AI tools were helpful during the brainstorming and revision stages because they provided useful suggestions and examples that made writing tasks easier to complete.

These findings are consistent with previous studies discussed in the literature review. Kessler (2018) stated that AI-generated feedback helps learners identify and correct writing errors more effectively. Similarly, Hwang et al. (2023) and Seufert et al. (2021) found that AI-supported feedback systems can support students in revising and organizing their writing more efficiently. The findings of the current study also support Kukulska-Hulme and Shield (2008), who argued that technology-supported learning environments can encourage learner autonomy and independent learning.

5.2 Concerns about overreliance on AI tools

Although students recognized many benefits of AI tools, the findings also revealed several concerns related to overreliance on technology. Many participants agreed that excessive use of AI tools could reduce regular writing practice, weaken creativity, and influence learning autonomy. Some interview participants also mentioned that they felt less confident writing independently without AI support.

These findings are similar to previous studies on the negative effects of AI-assisted writing. Zhai (2022) argued that overreliance on AI tools may reduce learners' critical thinking ability and independent writing skills. Likewise, Ranalli (2018) suggested that students who depend too heavily on automated feedback systems may become less confident in evaluating and improving their own writing independently.

Overall, the findings suggest that AI tools can provide useful support for writing development when they are used appropriately. However, students should use these technologies in a balanced way in order to avoid becoming too dependent on AI during the writing process.

6. Conclusion

The study investigated the pedagogical impact of AI tools on the writing skills of third-year English majors at HUFLIT. Based on the findings, the study showed that AI tools are commonly used by students in English writing activities. Many participants reported using applications such as ChatGPT, Gemini, Grammarly, and QuillBot to support different stages of the writing process, including brainstorming ideas, correcting grammar mistakes, improving vocabulary, and revising drafts.

The findings also indicated that most students had positive perceptions toward AI-assisted writing. Both the quantitative and qualitative results showed that students believed AI tools helped improve grammar accuracy, vocabulary use, coherence, organization, and idea development. In addition, many participants stated that AI tools helped them save time and complete writing tasks more effectively. The interview findings also revealed that students considered AI tools helpful for generating ideas and improving the clarity of their writing.

However, the study also identified several concerns related to the overuse of AI tools. Many students believed that depending too much on AI could reduce regular writing practice, weaken creativity, and affect independent learning. Some participants also mentioned that they felt less confident writing without AI support. These findings suggest that although AI tools provide many advantages, students still need to use them carefully and in a balanced way.

Overall, the study suggests that AI tools can positively support the writing development of English-major students when they are used appropriately. AI tools should be viewed as supportive learning resources rather than replacements for students' own thinking and writing ability.

Based on the findings, several recommendations can be made. First, teachers should provide students with guidance on how to use AI tools effectively and responsibly in writing activities. Students should be encouraged to use AI mainly for support, such as checking grammar, improving vocabulary, and developing ideas, instead of depending on AI to complete entire writing tasks. Second, students should continue practicing independent writing in order to strengthen their critical thinking, creativity, and confidence in writing. Finally, future research may investigate the impact of AI tools on other English language skills or involve students from different universities to provide broader understanding of AI-assisted language learning.

7. Limitations and Future Directions

This study has several limitations that should be acknowledged. The main limitation is that it did not directly measure students' actual writing performance. Instead, the findings were based on students' perceptions and self-reported experiences of using AI tools in writing. Therefore, the results should not be interpreted as evidence that AI tools objectively improve writing ability.

A second limitation is related to the use of self-report data. The questionnaire and interview responses may be affected by self-report bias, as participants relied on their own understanding and memory of their experiences. As a result, the reported perceptions may not fully represent their actual writing behavior.

Third, the study used convenience sampling and focused only on third-year English-major students at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). This sampling method may introduce selection bias because participation depended on students' availability and willingness to take part.

Another limitation is related to the qualitative phase. Although the interviews provided useful insights into students' experiences, only eight students were included. A larger and more diverse interview sample could have offered a broader range of perspectives on AI-assisted writing.

Based on these limitations, several directions for future research are suggested. First, future studies should include more objective measures of writing performance, such as writing tasks and analytic scoring rubrics, to examine whether students' perceived benefits of AI are reflected in their actual writing outcomes. Second, future research could compare students who use AI tools with those who do not, in order to better understand the specific impact of AI on the writing process. Longitudinal studies are also recommended to explore how students' perceptions, writing practices, confidence, and dependence on AI change over time. In addition, studies with participants from different universities, disciplines, and educational contexts would help improve the generalizability of the findings. Finally, future research may further investigate the relationships between AI use and factors such as learner autonomy, creativity, writing confidence, critical thinking, and AI literacy.

References

Cung, L. T., Hoang, U. P. T., Dinh, A. N., & Bui, T. H. (2025). University Students' perceptions of AI application in writing Skills in Vietnam: a Systematic review. Vietnam Journal of Education (Online)/Vietnam Journal of Education/GiáO DụC, 325–334. https://doi.org/10.52296/vje.2025.533

Hwang, Y., Lee, J. H., & Shin, D. (2023). What is prompt literacy? An exploratory study of language learners' development of new literacy skill using generative AI. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2311.05373

Kessler, G. (2018). Technology and the future of language teaching. Foreign Language Annals, 51(1), 205–218. https://doi.org/10.1111/flan.12318

Khan, A. L., Hasan, M. M., Islam, M. N., & Uddin, M. S. (2024). Artificial Intelligence Tools in Developing English Writing Skills: Bangladeshi University EFL Students' Perceptions. English Education Jurnal Tadris Bahasa Inggris, 17(2), 345–371. https://doi.org/10.24042/ee-jtbi.v17i2.24369

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. ReCALL, 20(3), 271–289. https://doi.org/10.1017/s0958344008000335

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. Heliyon, 7(5), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Ranalli, J. (2018). Automated written corrective feedback: how well can students make use of it? *Computer Assisted Language Learning*, *31*(7), 653–674. https://doi.org/10.1080/09588221.2018.1428994

Seufert, S., Guggemos, J., & Sailer, M. (2020). Technology-related knowledge, skills, and attitudes of pre- and in-service teachers: The current situation and emerging trends. Computers in Human Behavior, 115, 106552. https://doi.org/10.1016/j.chb.2020.106552

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. Frontiers in Psychology, 14, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Ta, T. T. H., Bo, T. L., Khuong, Q. N., & Nghiem, H. P. (2025). Impacts of artificial intelligence student guide on academic writing among third-year students at the faculty of international education, vnu university of languages and international studies. VNU Journal of Foreign Studies, 41(4), 40–53.https://doi.org/10.63023/2525-2445/jfs.ulis.5489

Tran, T. T. H. (2024). AI Tools in Teaching and Learning English Academic Writing Skills. Proceedings of the AsiaCALL International Conference, 4, 170–187. https://doi.org/10.54855/paic.23413

Utami, S. P. T., Andayani, A., Winarni, R., & Sumarwati, S. (2023). Utilization of artificial intelligence technology in an academic writing class: How do Indonesian students perceive? Contemporary Educational Technology, 15(4), ep450. https://doi.org/10.30935/cedtech/13419

Van, N. T. (2025). The impact of AI tools on English writing skills development. International Journal of Social Science Exceptional Research, 4(3), 01–06. https://doi.org/10.54660/ijsser.2025.4.3.01-06

Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. Journal of Research in Science Teaching, 59(10), 1765–1794. https://doi.org/10.1002/tea.21773

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*(1). https://doi.org/10.1186/s40561-024-00316-7

Bionote

Nguyen Minh Tuan holds a master's degree in TESOL and is currently an English lecturer at the Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). He has over 12 years of experience teaching English to both majoring and non-majoring students. His research interests include English language pedagogy, English for Specific Purposes (ESP), and the integration of technology in English language teaching (ELT).

Appendix A

Survey Questionnaire

Section I

Frequency of Using AI tools in writing skills

1. How often do you use AI tools in learning and completing your writing
- Never
- Rarely
- Sometimes
- Often
- Always

2. Which AI tool do you use in writing skill the most frequent?
- ChatGPT
- Gemini
- Grammarly
- Quillbot
- Others

Section II

Perceptions of Using AI tools in writing skills

Scale 1 2 3 4 5

Strongly Disagree Disagree Neutral Agree Strongly Agree

Item wording	Scale labels
AI tools are helpful in generating more ideas for writing.	1 2 3 4 5
AI tools provide clear and useful feedback.	1 2 3 4 5
AI tools help me avoid writing mistakes.	1 2 3 4 5
AI tools save my time when completing writing tasks.	1 2 3 4 5
AI tools challenge my writing effort.	1 2 3 4 5

Section III

Impacts of Using AI tools on writing developments

Item wording	Scale labels
AI tools help me improve my grammar accuracy.	1 2 3 4 5
AI tools help me use more appropriate vocabulary	1 2 3 4 5
AI tools improve the coherence and organization of my essays	1 2 3 4 5
AI tools help me develop ideas more clearly.	1 2 3 4 5
AI tools support me remember writing process.	1 2 3 4 5

Section IV

Concerns of Using AI tools in writing skill

Item wording	Scale labels
AI tools make me become dependent.	1 2 3 4 5
AI tools decrease my writing practice	1 2 3 4 5
AI tools increase the anxiety of making mistakes	1 2 3 4 5
AI tools influence my learning autonomy.	1 2 3 4 5
AI tools weaken my creativity	1 2 3 4 5

Appendix B

Interview Questions

1. Do you use AI tools when you write in English? Which ones?
2. How often do you use AI tools for writing tasks?
3. Do AI tools help you write better in English? How?
4. How do AI tools affect your writing skills overall?
5. Are there any problems when using AI tools for writing?
6. Will you continue using AI tools in the future? Why or why not?

#139 STUDENT EXPECTATIONS OF ENGLISH LANGUAGE TEACHERS: NAVIGATING THE SHIFT TO VIETNAM'S ESL TRANSITION AND AI INTEGRATION

Nguyễn Lê Uyên[1]

[1] *Nguyễn Lê Uyên, MA in TESOL, Faculty of Foreign Languages, HUFLIT, HCM City, Vietnam; Email: uyennl@huflit.edu.vn*

Abstract

In today's digital age, the amazing development of AI represents an unprecedented information revolution, permeating every profession especially English language teaching and learning. Alongside, Vietnam's transition of English from a foreign language to a second language has pushed a significant navigation in teaching and learning. These two simultaneous circumstances are profoundly changing the approach to English learning at all ages, particularly among Vietnamese students. Therefore, instructors at undergraduate level must re-evaluate their teaching methods and seek appropriate, engaging, and essential approaches for students' English language learning. This study aims to explore requirements and expectations of students regarding English language instructors' teaching methods in the era of AI and English becoming a second language in Vietnam. Based on a ten-question online survey using Google form carried out by 391 students majoring in English language from the Faculty of Foreign Languages at HUFLIT, its target is to identify the irreplaceable values and capabilities of instructors in English language teaching within the AI integration. The findings demonstrate students' preference for teachers as mentors, emotional supporters, cultural mediators, and soft-skill developers. As a result, instructors need to develop innovative methods to inspire learners, share learning pathways, and cultivate students' cultural as well as emotional intelligence when learning and using English.

Keywords**:** English Language Teaching, ESL Transition, AI Integration, Teacher Roles, Student Expectations

1. Introduction

We are living in an era where Artificial Intelligence (AI) is weaving itself into every aspect of life. Simultaneously, English Language Teaching (ELT) in Vietnam is shifting profoundly from a Foreign Language (EFL) to a Second Language (ESL) framework. Currently, we are in a transitional phase from EFL to ESL which is planning a nationwide implementation within several years. Policymakers in Vietnam recognized that English competence is a vital drive for international integration and economy growth. Therefore, the status of English has been altered officially and become an institutionalized medium of communication, collaboration, and global citizenship.

Inevitably, the explosion of AI tools has fundamentally changed traditional English language learning behaviors. University students today are digital natives who no longer rely solely on classroom teachers to access linguistic data, practice pronunciation, or verify grammatical accuracy. Free AI applications can provide personalized writing feedback, instant translation, and simulated conversation practice, so learners possess an unprecedented degree of language learning autonomy. Consequently, the non-stop rise of AI and the transition of Vietnam's emerging ESL landscape and has created an uncharted territory and immediate challenges for higher education. In an ESL-oriented classroom, students require high-level communicative competencies, critical thinking, and professional soft skills rather than only knowledge providing. The traditional and lecture-heavy pedagogical model is losing its efficacy to today learners when AI can perform basic linguistic corrections with instant response and flawless patience. To navigate this shift successfully, educators must understand the evolving expectations and changing psychology of students. Severe disconnect between classroom instruction and the actual needs of the future workforce may be created if educators fail to adapt curriculum design and teaching methodologies aligning with student realities.

While previous studies have examined the AI usage, teacher roles, and learner autonomy separately, this study explores the changing division of labor between AI tools and human teachers from students' perspective by examining what they expect from teachers under the combined pressure of AI integration and ESL transition. Therefore, this study is to investigate university student expectations regarding their English language teachers amidst the concurrent navigation toward an ESL framework and AI integration in Vietnam. To achieve its overall goal, the study is organized to answer three research sub-questions:

- What specific roles do students expect from their English teachers in the current AI-driven environment?
- To what extent do students believe AI tools should be integrated into English language pedagogy as Vietnam is shifting toward an ESL system?
- What human-led teaching qualities do students perceive as irreplaceable by AI in their language learning journey?

2. Literature Review

2.1. The shift from traditional to modern roles of English language teachers

2.1.1. From knowledge provider to strategic mentor

According to King (1993), the traditional educators function as ***"a sage on the stage"*** with the primary repository of raw information. As higher education environments evolve, the role of an instructor must change to foster critical engagement which is ***"a guide on the side"*** who actively structures spaces for collaborative knowledge construction. Moreover, as digital platforms decentralize how structural knowledge is accessed the primary value proposition of university faculty is forced to fundamentally reorganize from basic content delivery to high-level pedagogical support and structural instruction (Zawacki-Richter et al., 2019). Additionally, Nguyen (2021) highlights that within the regional context of Vietnam the transition away from teacher-centered instructional legacies requires a deep reconfiguration of traditional teacher identities as the absolute knowledge authority.

2.1.2. From task-bound compliance to competency-based innovation

Kluckhohn and Strodtbeck (1961) establish that through the framework of value orientations, legacy language classrooms operate on a ***"Doing" orientation*** that maximizes immediate, structured task actions over developmental progression. However, according to Hofstede and Hofstede (2005), while short-term orientation frameworks rely heavily on universal guidelines of right versus wrong, long-term oriented environments prioritize situational adaptability and structural perseverance. Therefore, classroom structures must adapt to lower power distance expectations, allowing flat-hierarchy consultation to replace top-down compliance when an educational system transitions toward active real-world utility. Similarly, Kluckhohn & Strodtbeck (1961) note that to align with contemporary competency goals, the pedagogical focus must pivot toward a ***"Growing"' orientation*** which prioritizes the internal intellectual evolution of each learner. Beckett & Miller (2006) add that by integrating professional simulations and media creation, innovative educators shift learner expectations from basic passive comprehension to institutionalized workforce competencies. The changing framework of English teaching is inline with results from previous studies, so this research will explore the exact applications in Vietnam's situation.

2.2. How AI should be integrated into English language pedagogy

2.2.1. In class: a co-facilitator for real-time critical thinking and collaboration

Huang et al. (2023) highlight in their analysis that AI integration forces a move away from assessing simple linguistic output to evaluating high-level cognitive processes. They stress that AI applications push language education toward fostering high-order thinking skills, critical evaluation, and creativity rather than simply testing memory or basic language mechanics. They also advocate for treating AI as a partner in the learning process, and pedagogy should focus on collaborative problem-solving and prompt-engineering literacy, where students learn to co-create and critically analyze AI-generated content.

Chassignol et al. (2018) emphasize that AI's primary function in the classroom environment is helping teachers save time in routine content delivery and grading and focus on more complex, interactive, and personalized human coaching. Furthermore, Chan and Colloton (2024) argue higher education must adapt to the presence of AI by reshaping in-class hours into active workshops, more complex, interactive, and personalized human coaching. That is the similar scenario of Vietnamese classroom orientation in higher education.

2.2.2. At home: a personalized "linguistic sandbox"

According to Koraishi (2023), for students who are too intimidated to speak in class, interacting with AI provides a low-stakes environment because it offers a non-judgmental space for language production, effectively mitigating foreign language speaking anxiety by removing the fear of negative social evaluation by peers or instructors. In addition, Stockwell (2021) defines the ideal out-of-class digital ecosystem as one that extends learning beyond the physical boundaries of the institution, empowering learners to manage their own linguistic remediation and practice schedules autonomously to build student independence and prepare for real-life interaction. Concurrently, Godwin-Jones (2024) and Zhai (2023) discuss using AI language models at home to manage the heavy lifting of grammar drilling, spell-checking, and foundational editing. This is a fact that Vietnamese students are rushing for using AI to improve English language themselves at la ow cost and high comfort.

2.3. Irreplaceable qualities of human-led teaching in English language pedagogy

2.3.1. Emotional Intelligence (EQ) and relational encouragement

According to Dörnyei & Ushioda (2011), as language learning requires long-term psychological persistence, a teacher's role cannot be reduced to only transmission of content but an active source of motivational support which deeply depends on the interpersonal rapport and authentic encouragement generated between the teacher and the learner. Furthermore, automated tools may correct structural errors, but they cannot replace the emotional support and interpersonal rapport required by learners (Mercer & Dörnyei, 2020). Consequently, Mercer and Dörnyei (2020) emphasize the contemporary language pedagogy must prioritize the socio-emotional well-being and active engagement of the learner - a position that requires educators center on creating a safe, humanistic space for students' growth. Similarly, Vietnamese students are increasingly showing their needs for teachers as mentors and emotional supporters.

2.3.2. Cultural bridge and social intelligence in communication

The theoretical model of Vygotsky (1978) implies that linguistic competencies are co-constructed within a cultural context which means dialogue with a human mentor is structurally superior to isolated interactions with static machine intelligence. Furthermore, Byram (1997) establishes that authentic intercultural communicative competence requires learners to develop specific cultural attitudes, critical awareness, and the social intelligence to navigate diverse human relationships. He also concentrates on the must of instructors serving as a cultural

bridge, guiding students through the sociolinguistic complexities that automated translation models ignore. In this context, effective intercultural communication also requires learners to interpret meaning appropriately across social and cultural contexts, including situational appropriateness and non-verbal cues (Byram, 1997). Such social intelligence cannot be developed from statistical text models alone because appropriate communicative choices are bound to real-world social consequences and human relationships, which requires a human mentor possessing authentic social intelligence and lived cultural experience.

3. Methodology

A mixed-method (mainly quantitative survey with one qualitative open-ended item) was adopted in this study to capture student expectations, mapping out how learners perceive the boundary between machine-driven automation and human-led mentorship amidst emerging AI integration and the shifting ESL landscape in Vietnam. The nine closed questions, using a 5-point Likert scale or multiple-choice grids, and one open-ended question were distributed through online communication. The instrument development relies on three sub-research questions and a literature framework. In more detail, items 1 to 3 explore expected teacher roles, while questions 4 to 6 show the extension of applying AI in students' learning, and the third examination of irreplaceable human teachers' qualities is covered by numbers 7 to 9. The last question, which is open ended, pertains to students' concept of "*a perfect English language teacher*", allowing participants to elaborate freely in their own words. It provides unconstrained qualitative depth in their expectations for future English language teachers. The questionnaires were checked once again by the author's colleague before being sent to participants.

The population for this study comprises 391 undergraduate English majors at the Faculty of Foreign Languages at HUFLIT (Ho Chi Minh City University of Foreign Languages and Information Technology), Vietnam. They are students from the first to final years in their undergraduate program who have learned with this paper's author in some courses, and they were recruited voluntarily online. They have all gone through the English language foundation stage and are fluent users of AI in their learning.

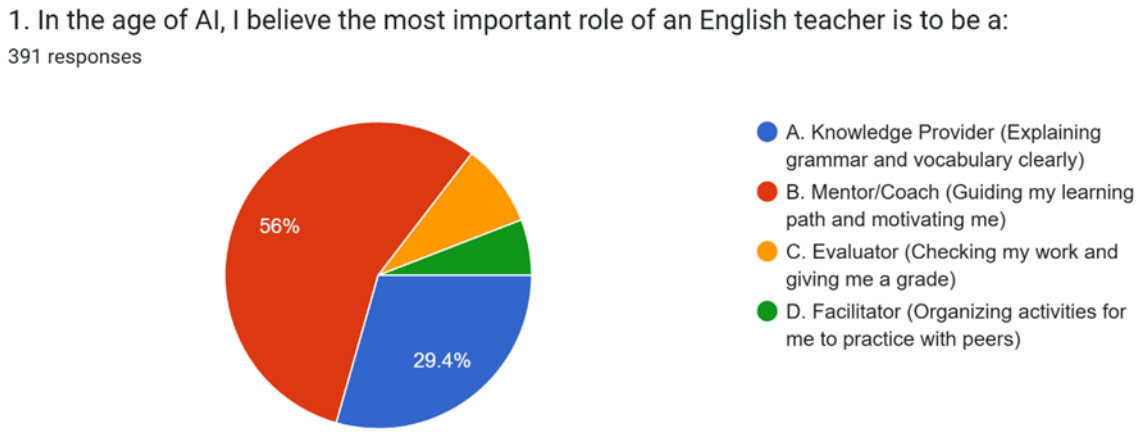
The data collection was bounded within a one-week period (in May 2026) to ensure responses were anonymous and consistent in maintaining objectivity. The collected data were extracted into a centralized spreadsheet and then cleaned to remove incomplete or duplicate entries. Each item was analyzed independently and later synthesized to address the three sub-research questions. For item number 10, data were subjected to iterative cycles of open coding; similar phrases were clustered into codes, which were synthesized into themes to explain the reasons behind quantitative trends.

4. Results

This section includes findings of a survey carried out online by 391 English language students at Faculty of Foreign Languages at HUFLIT, Vietnam.

Figure 1

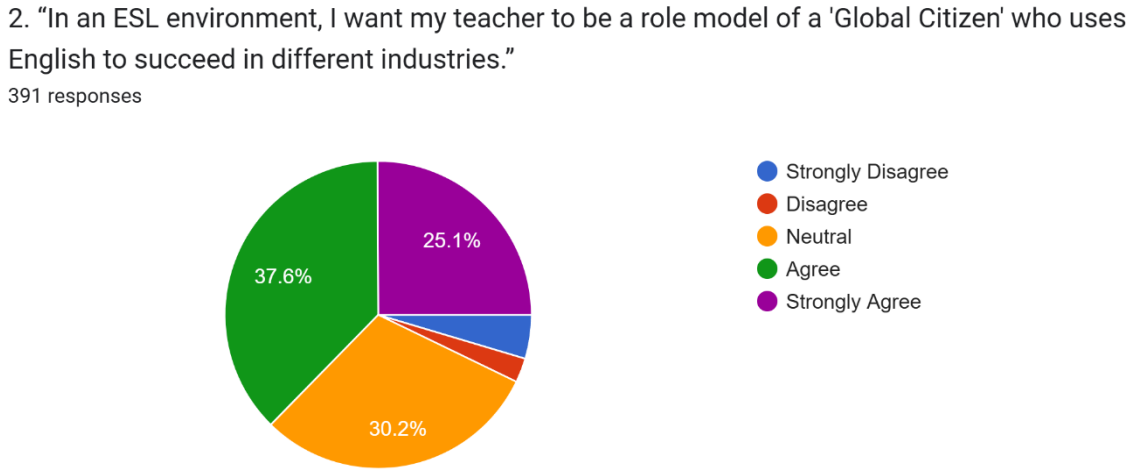
The most important role of an English language teacher



The findings reveal a noticeable shift in students' expectations of teachers, moving from traditional teacher-centered roles toward more supportive and learner-centered functions. Most respondents (56%) identified the role of mentor or coach as the most important, suggesting that students increasingly expect teachers to provide guidance, motivation, personalized support, and assistance in navigating their learning journey. The role of knowledge provider ranked second, accounting for 29.4% of responses. Although this proportion is considerably lower than that of the mentor/coach role, it still represents nearly one-third of the sample, indicating that students continue to value teachers' expertise in explaining complex language concepts, correcting misunderstandings, and providing professional insights that AI may not always deliver effectively. In contrast, the roles of evaluator and facilitator received much lower levels of support, with approximately 8.7% and 5.9% of respondents selecting these options, respectively. In brief, this pie chart provides similar findings comparing with that of King (1993). The data suggests that students are actively pushing away from the traditional, transmissive ***"sage on the stage"*** (the 29.4% legacy expectation) and demanding a ***"guide on the side"*** who acts as a strategic mentor (56%).

Figure 2

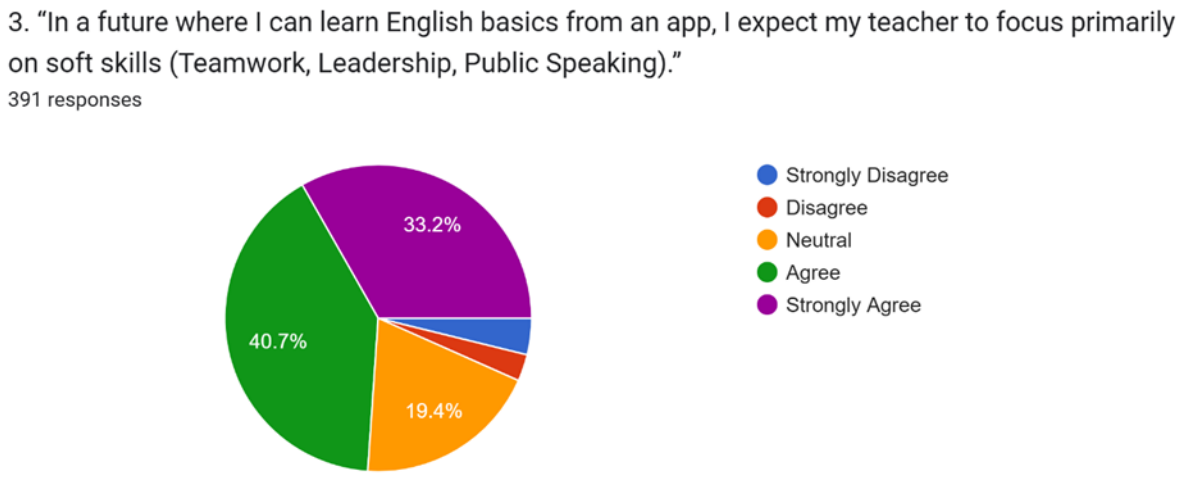
The "Global Citizen" role of an English language teacher in an ESL environment



As clearly shown, the largest proportion of respondents selected "Agree" (37.6%), while a further 25.1% chose "Strongly Agree," resulting in a combined positive response rate of 62.7%. This indicates that nearly two-thirds of the participants expect English language teachers not only to possess strong language proficiency but also to serve as role models who can illustrate how English facilitates success in international communication, career development, and cross-cultural interactions. Meanwhile, 30.2% of respondents remained neutral, representing a substantial proportion of the sample. In contrast, only a small minority expressed negative views, with approximately 3.3% disagreeing and 3.8% strongly disagreeing, accounting for a combined total of 7.1% of responses. In conclusion, this strong positive shift (62.7% combined agreement) directly follow Byram's (1997) ***Intercultural Communicative Competence*** framework. As Vietnam navigates its emerging ESL transition, students look to their lecturers not just for local text comprehension, but to act as lived cultural mediators who model how to negotiate diverse professional industries globally.

Figure 3

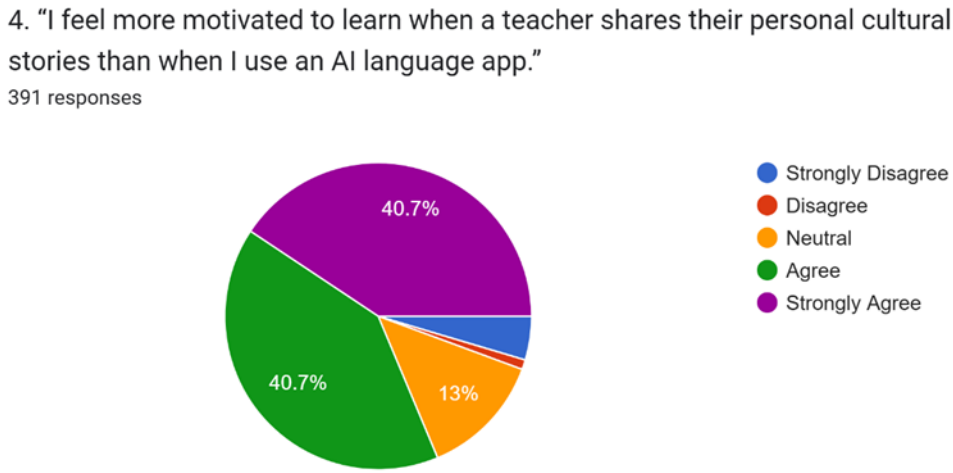
The students' need for soft skills practicing in class



The largest proportion of respondents selected "Agree" (40.7%), while a further 33.2% chose "Strongly Agree," resulting in a combined agreement rate of 73.9%. This overwhelming majority indicates that most learners expect teachers to devote greater attention to fostering soft skills that cannot be easily acquired through language-learning applications alone. Meanwhile, 19.4% of respondents selected the neutral option, suggesting that although many students support this shift in teaching priorities, a notable proportion remain uncertain about the extent to which soft skills should be integrated into English language education. In contrast, only a small minority expressed disagreement, with approximately 3.3% disagreeing and 3.4% strongly disagreeing, accounting for less than 7% of the total responses. In general, this is a shift from a task-bound ***"Doing" orientation*** to a ***"Growing" orientation*** (Kluckhohn & Strodtbeck, 1961). Because automated apps can efficiently triage linguistic baselines asynchronously at home (Godwin-Jones, 2024), students need valuable in-class human hours dedicated to socio-adaptive human competencies.

Figure 4

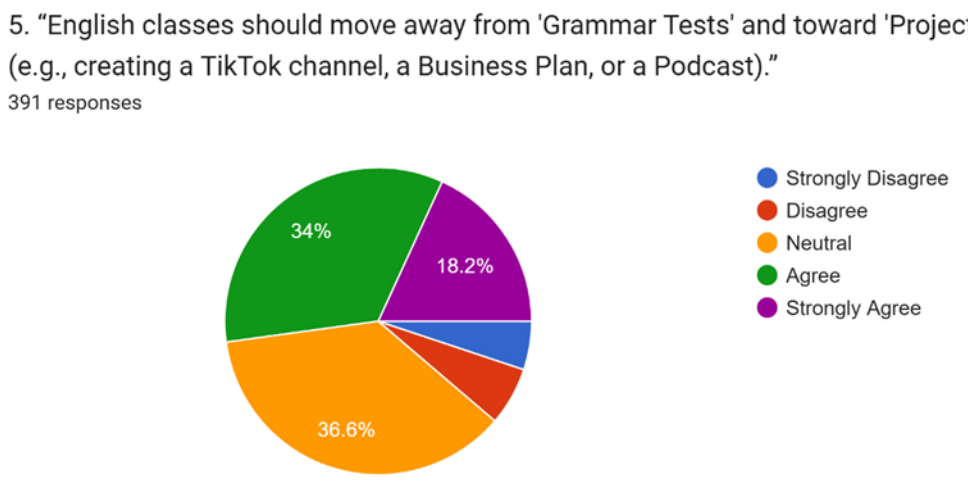
The importance of sharing teacher's personal cultural experiences



The findings indicate a very strong preference for human interaction and authentic cultural engagement in the language learning process. A majority of respondents expressed positive attitudes toward the statement, with 40.7% selecting "Agree" and another 40.7% choosing "Strongly Agree." Combined, these categories account for 81.4% of all responses, demonstrating that a substantial majority of learners find teachers' personal experiences and cultural narratives more motivating than the use of AI-based language learning applications. Meanwhile, 13% of respondents selected the neutral option, indicating that some learners were uncertain or believed that both teachers and AI applications can play complementary roles in supporting motivation. In contrast, only a very small minority expressed negative views, with approximately 1.0% disagreeing and 4.6% strongly disagreeing. All these results are consistent with Dörnyei and Ushioda's (2011) ***model of relational motivation***. It proves that long-term second language learning stamina cannot be sustained by cold data transfer; it is a dynamic process sparked by real human chemistry, vulnerability, and shared personal narratives.

Figure 5

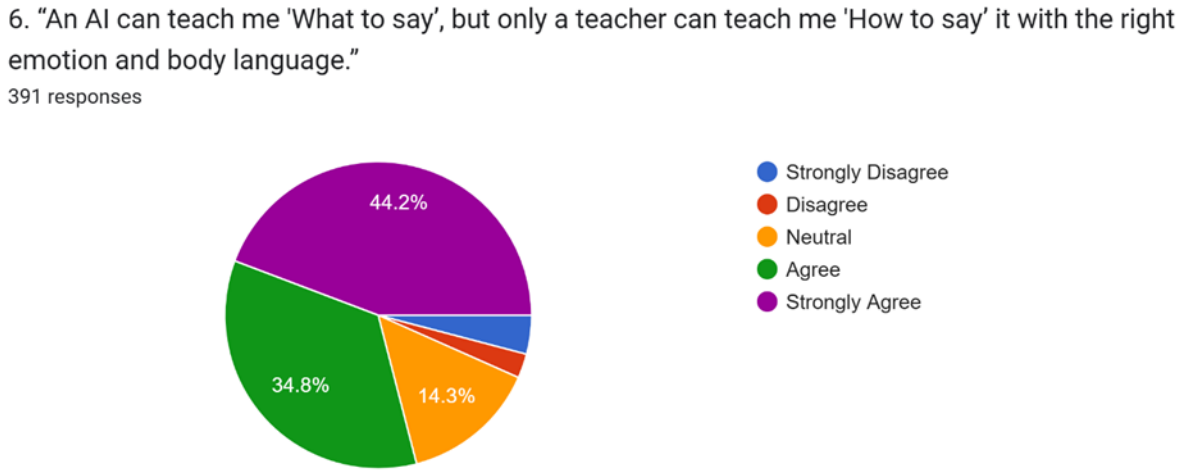
The preference for Project-based Learning to Grammar Tests



The largest proportion of respondents selected the neutral option (36.6%), indicating that a considerable number of students remain undecided about replacing traditional grammar-focused assessments with project-based activities. Nevertheless, positive responses outweighed negative ones, with 34% of participants agreeing and 18.2% strongly agreeing with the statement. Combined, these categories account for 52.2% of all responses, suggesting that more than half of the respondents support a shift toward more practical, creative, and application-oriented learning experiences. In contrast, only a relatively small proportion of participants expressed disagreement, with approximately 6.9% selecting "Disagree" and 4.3% choosing "Strongly Disagree," resulting in a combined negative response rate of 11.2%. In general, the 52.2% combined agreement aligns directly with Beckett and Miller (2006) regarding ***Project-Based Second Language Education***. The heavy neutral chunk (36.6%), however, likely reflects the systemic friction noted by Nguyen (2021): changing institutional testing habits from short-term memorization (Hofstede's STO) to authentic task outputs can trigger learner uncertainty when students are highly conditioned by traditional exam structures.

Figure 6

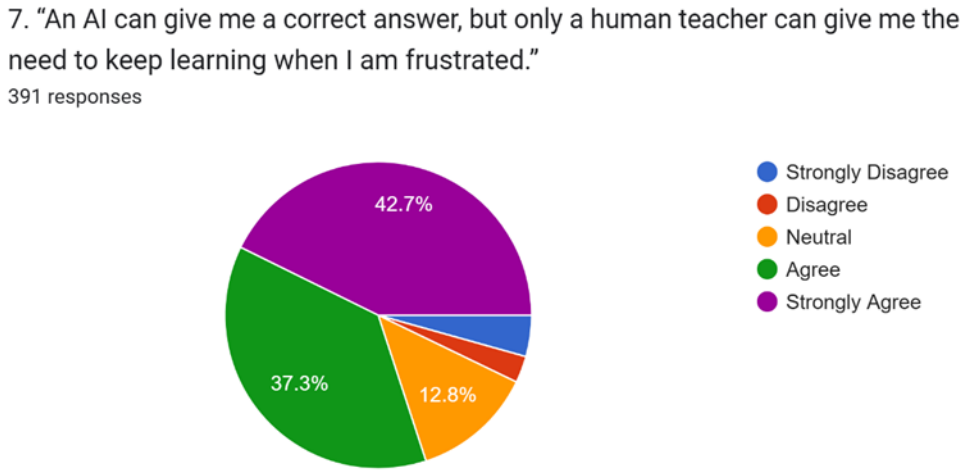
The preference for learning "how to say" to "what to say"



The highest percentage of responses fell into the "Strongly Agree" category, accounting for 44.2% of the total survey sample, which was closely followed by the "Agree" cohort at 34.8%. Combined, these two figures represent a resounding 79.0% of the total respondents who acknowledged that while AI can sufficiently guide learners on factual content (what to say), human teachers remain vital for mastering delivery skills (how to say). In stark contrast, only 14.3% of the surveyed individuals maintained a "Neutral" stance on the matter. The remaining segments represented the dissenting opinions, which altogether comprised less than 8% of the data; specifically, those who expressed "Strongly Disagree" and "Disagree" accounted for approximately 4.3% and 2.4% respectively, highlighting a minimal resistance among the sample to the premise of the statement. In sum, this item's result reinforces Byram's (1997) emphasis on the social and cultural dimensions of communicative competence. While advanced large language models are exceptional at rectifying syntactic errors ("what to say"), students recognize that the situational appropriateness and non-verbal subtext ("how to say it") require a human mentor with lived social experience.

Figure 7

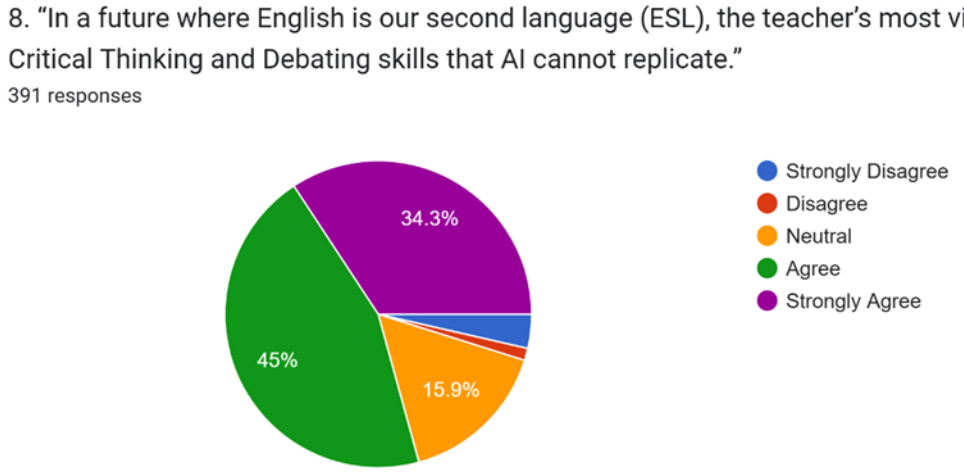
The importance of Encouragement from an English language teacher



To break down the figures, 42.7% of the respondents expressed the highest level of conviction by selecting "Strongly Agree", while a substantial 37.3% chose "Agree". Together, this cumulative 80.0% dominates the chart, validating the hypothesis that technical correctness alone cannot replace a teacher's emotional support. On the other hand, the proportion of individuals remaining "Neutral" was limited to 12.8%. The remaining minority who questioned the statement accounted for less than one-tenth of the total sample, with 4.3% of respondents stating they "Strongly Disagree" and a mere 2.9% opting for "Disagree", demonstrating that opposition to the psychological necessity of human educators is statistically minimal. Overall, this is in line with Mercer and Dörnyei's (2020) emphasis on the socio-emotional dimensions of language learning. Language acquisition is emotionally volatile; when students experience cognitive blocks due to anxiety, they require the high emotional intelligence and empathy of a human mentor (Mercer & Dörnyei, 2020) to lower their affective filters.

Figure 8

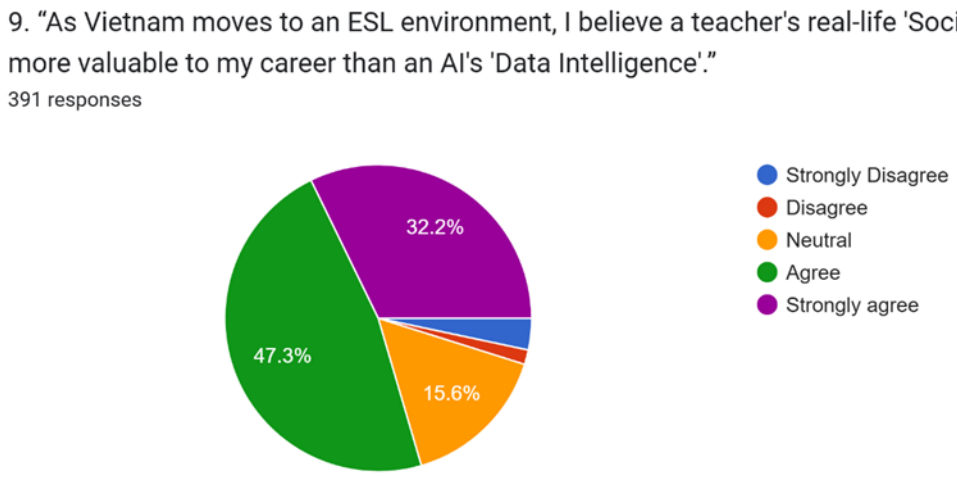
The importance of practicing Critical Thinking and debating skills in class



The “Agree” category emerged as the largest segment, accounting for exactly 45% of the total distribution, while the “Strongly Agree” cohort constituted 34.3%. Together, these two figures represent a substantial cumulative majority of 79.3%, indicating a robust belief that pedagogical priorities must evolve toward skills that surpass automated capabilities. Conversely, the proportion of individuals who maintained a “Neutral” attitude stood at 15.9%. Dissenting viewpoints were minimal, collectively representing less than 5% of the survey population; specifically, the “Strongly Disagree” and “Disagree” sectors accounted for minor fractions of approximately 3.5% and 1.3% respectively, confirming that resistance to this predicted shift in ESL instruction is statistically negligible. In brief, this result is consistent with Huang et al. (2023). Since generative technology can instantaneously fabricate a grammatically flawless essay or script, the instructor's modern role must shift away from checking low-level mechanics to fostering high-order critical evaluation, and active dialectic debate.

Figure 9

The comparison between Social Intelligence and Data Intelligence



It is highly apparent that a substantial majority of the surveyed individuals expressed a positive attitude toward this notion, reinforcing the perceived superiority of human interpersonal skills over algorithmic data processing in a developing professional landscape, while neutral and negative stances remained in the minority. In detail, the “Agree” cohort constituted the largest share of the dataset, representing nearly half of the total distribution at 47.3%. This was supported by an additional 32.2% of participants who registered a “Strongly Agree” response, bringing the cumulative positive consensus to a significant 79.5%. In contrast, the proportion of individuals who maintained a 'Neutral' position was limited to 15.6%. Active opposition to the statement was extremely minimal, accounting for less than 5% of the total sample; specifically, those who chose “Strongly Disagree” made up approximately 3.5%, while the “Disagree” category recorded the lowest frequency at a mere 1.4%, indicating that skepticism regarding the premium placed on human social intelligence is statistically negligible. In general, this finding serves as a firm support to Vygotsky’s (1978) ***Social Constructivism.*** Students are declaring that as language is a social tool meant to navigate real human relationships in the workforce, learning it requires a co-facilitated, socially intelligent environment. They gracefully accept AI’s “data intelligence” as a supportive backend, but guard the upper-level pedagogical space for human-to-human interaction.

Figure 10

The qualities of a "Perfect English Teacher"' of the year 2030

Core Pedagogical Theme / Phrase	Exemplary Student Phrase	Frequency of Repetition (N=391)
Communication Skills over Grammar	*"Helps every student become a confident communicator instead of only teaching grammar rules."*	**117**
Personalized Learning & Individual Needs	*"Focuses more on students' individual needs and personalized teaching methods."*	**94**
Creativity & Classroom Engagement	*"Uses interactive games, technology, and creative projects to make learning interesting."*	**73**
Critical Thinking Facilitator	*"Guides students to improve critical thinking, evaluation, and problem-solving."*	**70**
Real-life & Workforce Application	*"Focuses on real-life English skills, career readiness, and practical applications."*	**70**
Strategic Mentor & Guidance	*"Acts as a mentor who gives guidance rather than just acting as a traditional knowledge provider."*	**63**
Source of Motivation & Inspiration	*"An inspiring presence that motivates students to achieve their long-term language goals."*	**62**
Empathy & High Emotional Intelligence (EQ)	*"Possesses high emotional intelligence to understand student frustration and provide psychological encouragement."*	**39**
Holistic Soft Skills Developer	*"Uses class time to automate mechanical grading and focus entirely on developing soft skills."*	**19**

By coding and counting similar answers (using Excel software) and grouping them into themes with exemplary phrases, the spontaneous and wide range expectation of students about *"perfect English language teachers"* are simplified into the above table. It can be explored from the table that the highest repetition belongs to ***Communication Skills*** (117 counts). This suggests that students recognize the official shift toward an ESL environment where functional, real-world spoken fluency vastly outweighs legacy, task-bound grammar drilling. About the ***"Guide on the Side"*** validation**, Personalized Learning** (94), **Critical Thinking** (70), and **Mentorship** (63) score highly which proves an alignment with King's (1993) framework. Students are explicitly asking for teachers to abandon the "Knowledge Provider" monopoly and step into the role of a strategic coach. Concerning ***the Irreplaceable Human Factor,*** even though technology is mentioned, the demand for ***Motivation*** (62) and ***Empathy/EQ*** (39) show that students openly state a teacher's value lies in providing the safe psychological space and encouragement which an AI platform cannot simulate.

5. Discussion

The findings of this study demonstrate a profound shift in undergraduate expectations of English language lecturers within the Vietnamese higher education context. As advanced

digital systems democratize baseline language mechanics, students are actively re-evaluating the instructional division of labor between human educators and artificial intelligence.

5.1. The operational shift from knowledge provider to strategic mentor

The quantitative data validate the theoretical transition from a traditional ***"sage on the stage"*** to a modern ***"guide on the side"*** (King, 1993). When asked to define the primary role of an English teacher in the age of AI (Item 1), a decisive majority (56%) selected ***"Mentor/Coach"***, while only 29.4% favored the legacy role of ***"Knowledge Provider"***. This structural pivot is corroborated by Item 3, where 73.9% of respondents (Agree and Strongly Agree combined) expect instructors to focus primarily on high-level soft skills like teamwork, leadership, and public speaking rather than language basics. However, the neutral group may suggest that students are interested in innovation but still uncertain about replacing familiar assessment practices and the traditional role of teachers as a knowledge provider. Viewed through the value orientation framework of Kluckhohn and Strodtbeck (1961), these percentages indicate an institutional shift from a task-bound ***"Doing" orientation*** (focused on immediate test compliance) to a student-centric ***"Growing" orientation*** (focused on long-term capability). Students openly expect human-led class hours to be liberated from mechanical instruction, a demand further reflected in the qualitative data where ***Personalized Learning & Individual Needs*** (n = 94) and ***Strategic Mentorship*** (n = 63) emerged as dominant criteria for the future "Perfect Teacher".

5.2. Curricular Innovation: Navigating the Emerging ESL Landscape

As Vietnam actively transitions from an EFL toward a formalized ESL framework to cultivate global citizenship, student demands for structural curricular design have evolved. A significant majority of participants (52.2%) expressed agreement or strong agreement that language classes should move away from standard grammar examinations toward complex project-based Learning (item 5), such as producing podcasts or business plans. Furthermore, 79.3% of undergraduates agree that in an ESL ecosystem, the teacher's most vital job is to cultivate critical thinking and debating skills that algorithms cannot replicate (Item 8). This orientation aligns clearly with Hofstede's (2005) conceptualization of a ***Long-term Orientation***, wherein educational success is measured by situational adaptability and real-world utility rather than short-term rote recall. The evidence proves that students no longer view the classroom as an isolated syntactic training ground, but rather as an incubator for professional readiness, which is echoed in the qualitative findings ranking ***Communication Skills over Grammar*** (n = 117) and ***Real-life & Workforce Application*** (n = 70) as the highest student priorities.

5.3. The Irreplaceable Human Core: Emotional and Social Intelligence

The most impressive data emerges around the affective and socio-cultural dimensions of language acquisition, exposing the strict limitations of artificial intelligence. While AI tools are recognized for their data intelligence, students display a massive preference for human interaction when navigating the vulnerable process of learning a language. Specifically, 81.4% of students feel more motivated when an instructor shares personal cultural experiences and lived narratives than when interacting with an AI language application (Item 4). 80% agree that while AI can provide accurate technical answers, only a human teacher can deliver the authentic ***"Encouragement"*** needed to overcome learning frustration (Item 7). This data provides support for Dörnyei and Ushioda's (2011) emphasis on motivation and Mercer and Dörnyei's (2020) emphasis on socio-emotional engagement, suggesting that sustained linguistic stamina requires genuine human empathy and psychological validation.

Furthermore, communication is recognized as a deeply contextual, symbolic phenomenon rather than a simple transactional data transfer. This is evidenced by Item 6, where a striking 79% of participants agreed that *"An AI can teach me 'What to say', but only a teacher can teach me 'How to say' it with the right emotion and body language."* This realization is reinforced by Item 9, where 79.5% of respondents asserted that a human teacher's real-world ***"Social Intelligence"*** is vastly more valuable to their careers than an AI's ***"Data Intelligence."***

Consequently, the human teacher remains entirely irreplaceable as a lived cultural bridge (Byram, 1997) and a model for intercultural communicative competence (Byram, 1997) – which is reinforced by the qualitative clusters identifying ***Empathy and High EQ*** (n = 39) and ***Social Intelligence*** as core human pillars.

6. Conclusion

6.1. Empirical and theoretical summary

This study maps the evolving expectations of N = 391 undergraduate English majors in Vietnam against the dual backdrops of rapid artificial intelligence integration and an official transition toward an ESL educational framework. The data demonstrate a structural consensus: modern language learners do not view AI as a replacement for human educators, but rather as a tool to catalyze a long-awaited pedagogical evolution. Students explicitly relegate lower-level, mechanical linguistic triage - such as basic grammar explanation and automated text corrections - to machine datasets, viewing out-of-class AI interactions as a safe, low-stakes "linguistic sandbox" to lower foreign language anxiety. Concurrently, this spatial division of labor completely redefines the institutional value proposition of the human lecturer. The physical university classroom is no longer expected to operate as a transmission hub for rote data delivery. Instead, students decisively demand that human educators step into roles as strategic mentors, cultural mediators, and coaches of advanced soft skills. The overwhelming majority of participants emphasize that while an algorithm operates on "data intelligence", it remains decoupled from the affective, socio-cultural, and relational realities of language acquisition. True linguistic fluency, pragmatic delivery, and professional confidence are co-constructed through human interaction, relying on a teacher's real-world "social intelligence", empathy, and relational encouragement. Ultimately, the "Perfect English Teacher" of the future is not defined by their mastery over syntactic rules, but by their high emotional intelligence and capacity to guide students toward becoming global citizens.

6.2. Limitations of the study

While this study provides robust data-driven insights, several structural limitations must be acknowledged:

- **Geographic and institutional sample isolation**

The sample size (N = 391) was strictly recruited from a single private university in Ho Chi Minh City via convenience sampling. Consequently, the findings may not fully represent student expectations in public universities or institutions located in rural and suburban regions of Vietnam, where resource access and baseline digital literacy may differ.

- **Static cross-sectional design**

The data capture student expectations at a single point in time. Because generative artificial intelligence technologies and national ESL policies are evolving at an accelerated pace, student attitudes and reliance on digital tools may shift dynamically over a longer timeline.

- **Self-reported data bias**

The study relies heavily on self-reported survey perceptions. While useful for mapping internal attitudes, student expectations of what they want from a curriculum do not always directly correlate with objective measures of learning efficiency or actual academic performance.

- **Instrumental insufficiency**

The questionnaire is short and developed by the author basing on a simple framework of three sub-research questions, so it lacks full reliability and validity. The study also relies on one open-ended question for qualitative data, which limits the depth of interpretation. Another concern is that the findings capture broad student expectations rather than detailed explanations of why students hold these views.

6.3. Recommendations for future research

To build upon the empirical foundations established by this study, future research should explore the following avenues:

- **Longitudinal inquiries**

Future studies should deploy longitudinal research designs to track how undergraduate expectations, prompt literacy, and AI dependencies evolve over their four-year university trajectory as generative tools become deeply normalized in secondary education.

- **Comparative institutional analyses**

Expanding data collection to include a comparative analysis between private and public universities, as well as northern, central, and southern regional institutions, would provide a more generalized national mapping of the ESL transition.

- **Multi-perspective triangulation**

Future projects should actively capture faculty perspectives, exploring the institutional challenges, anxieties, and training deficiencies experienced by university lecturers as they attempt to transition from traditional “knowledge providers” to “strategic mentors”. Moreover, future research should include interviews, focus groups, or classroom observation to triangulate survey findings.

In brief, future studies could compare student and teacher expectations, examine whether AI literacy changes students’ expectations over time, or test whether redesigned teacher roles actually improve learner engagement, communication confidence, and critical thinking. This would move the research beyond perception mapping toward pedagogical evaluation.

References

Beckett, G. H., & Miller, P. C. (Eds.). (2006). Project-based second and foreign language education: Past, present, and future. Information Age Publishing. https://doi.org/10.1108/978-1-60752-716-9

Byram, M. (1997). Teaching and assessing intercultural communicative competence. Multilingual Matters.

Chan, C. K. Y., & Colloton, T. (2024). Generative AI in higher education. Routledge.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. Procedia Computer Science, 136, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Dörnyei, Z., & Ushioda, E. (2011). Teaching and researching motivation. Longman.

Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. Language Learning & Technology, 28(2), 5–31. https://doi.org/10.64152/10125/73570

Hofstede, G. H., & Hofstede, G. J. (2005). Cultures and organizations: Software of the mind (Rev. and expanded 2nd ed.). McGraw-Hill.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. Educational Technology & Society, 26(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).0009

King, A. (1993). From sage on the stage to guide on the side. College Teaching, 41(1), 30–35. https://doi.org/10.1080/87567555.1993.9926781

Kluckhohn, F. R., & Strodtbeck, F. L. (1961). Variations in value orientations. Row, Peterson and Company.

Koraishi, O. (2023). Exploratory study on the use of ChatGPT in alleviating foreign language speaking anxiety. Journal of Education and E-Learning Research, 10(3), 424–432.

Mercer, S., & Dörnyei, Z. (2020). Engaging language learners in contemporary classrooms. Cambridge University Press. https://doi.org/10.1017/9781009024563

Nguyen, T. M. H. (2021). Teacher professional development in Malaysia and Vietnam: Shifting paradigms in language teacher education. Journal of Asia TEFL, 18(3), 912–926.

Stockwell, G. (2021). Mobile assisted language learning: Concepts, contexts and challenges. Cambridge University Press. https://doi.org/10.1017/9781108652087

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2023). ChatGPT for next generation science learning. XRDS: Crossroads, The ACM Magazine for Students, 29(3), 42–46. https://doi.org/10.1145/3589649

Bionote

The author of this paper has been a lecturer at universities in Ho Chi Minh City in fifteen years, and sustains a deep interest in subjects such as language skills, translating, interpreting, literature, civilization, and intercultural communication.

Appendices

Questionnaires

Student Survey

The Future of English Learning in the AI & ESL Era

Section 1

The Teacher as a Personal Brand & Evolving Role

1. In the age of AI, I believe the most important role of an English teacher is to be a:

- Knowledge Provider (Explaining grammar and vocabulary clearly)
- Mentor/Coach (Guiding my learning path and motivating me)
- Evaluator (Checking my work and giving me a grade)
- Facilitator (Organizing activities for me to practice with peers)

2. In an ESL environment, I want my teacher to be a role model of a *'Global Citizen'* who uses English to succeed in different industries.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

3. In a future where I can learn English basics from an app, I expect my teacher to focus primarily on soft skills (Teamwork, Leadership, and Public Speaking, etc.)

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Section 2

Pedagogical Innovation & Practicality

4. I feel more motivated to learn when a teacher shares their personal cultural experiences and stories than when I use an AI language app.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

5. English classes should move away from *'Grammar Tests'* and toward *'Project-Based Learning*'.

- Strongly Disagree
- Disagree
- Neutral

- Agree
- Strongly Agree

6. An AI can teach me *'What to say'*, but only a teacher can teach me *'How to say'* it with the right emotion and body language.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Section 3

The Irreplaceable Human Element

7. An AI can give me a correct answer, but only a human teacher can give me the *'Encouragement'* I need to keep learning when I am frustrated.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

8. In a future where English is our second language (ESL), the teacher's most vital job is to teach Critical Thinking and Debating skills that AI cannot replicate.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

9. As Vietnam moves to an ESL environment, I believe a teacher's real-life *'Social Intelligence'* is more valuable to my career than an AI's *'Data Intelligence'*.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Section 4

The "Dream" Classroom

10. Describe the *'Perfect English Teacher'* of the year 2030. What do they do differently than teachers today?

#140 RETHINKING ASSESSMENT IN THE SHIFT FROM EFL TO ESL: A CRITICAL REVIEW OF NO-EXAM CLASSROOMS

Thi Thanh Thao Nguyen[1], Thi Dieu Duyen Le[2], Thi Hong Tham Chau[3], Thanh Tinh Bui[4], Van Doan Tran[5]

Abstract

This article critically examines the growing interest in "no-exam" classrooms within the transition from English as a Foreign Language (EFL) to English as a Second Language (ESL) in Vietnam. While traditional exam-based assessment is increasingly viewed as misaligned with communicative and context-sensitive language use, alternative approaches that emphasize continuous, evidence-based assessment are often presented as inherently superior. Drawing on perspectives from language testing, sociocultural theory, and contemporary assessment scholarship, this paper argues that the shift from exams to evidence does not resolve longstanding tensions in assessment but rather redistributes them. In particular, issues of validity, reliability, learner agency, and institutional accountability remain central. The paper proposes a reconceptualised assessment framework that integrates multimodal evidence, feedback processes, and evaluative judgement to navigate the inherent trade-offs associated with these constraints. It concludes that the future of TESOL assessment lies not in abandoning exams, but in developing hybrid systems that acknowledge and negotiate the tensions between pedagogical relevance and institutional credibility.

***Keywords*:** assessment reform, no-exam classroom, evaluative judgement, ESL transition, feedback processes

1. Introduction

The ongoing shift from English as a Foreign Language (EFL) to English as a Second Language (ESL) in Vietnam necessitates a fundamental reconsideration of how language learning is assessed. As educational priorities increasingly emphasize communicative competence, real-world language use, and functional multilingualism, traditional assessment systems, which have been largely dominated by high-stakes examinations, appear increasingly misaligned with these goals (Fulcher, 2010). While exams have historically been associated with objectivity, comparability, and institutional accountability, their capacity to capture the dynamic, socially situated, and developmental nature of language use remains contested (Messick, 1989; Fulcher, 2010).

This misalignment reflects a deeper epistemological tension between the nature of language and the logic of assessment. Language is not a static body of knowledge to be reproduced under time pressure, but a socially mediated and evolving capacity shaped through interaction, feedback, and meaningful communicative practice (Vygotsky, 1978). Yet, exam-based systems continue to conceptualize language as discrete, decontextualized knowledge, privileging accuracy and performance under constrained conditions. As a result, learners may succeed in examinations while remaining unable to use language effectively in authentic contexts.

[1] *Corresponding author, International Institute, Ho Chi Minh City University of Economics and Finance, Vietnam, Email: thaonnt2@uef.edu.vn*
[2] *Faculty of Foreign Languages & Oriental Studies, Yersin University of Da Lat, Vietnam*
[3] *Faculty of Foreign Language, Nguyen Tat Thanh University, Vietnam*
[4] *Foreign Languages and Informatics Center, Dong Thap University, Vietnam*
[5] *Faculty of Foreign Languages, Ho Chi Minh City University of Industry and Trade, Vietnam*

This tension is further reinforced by the institutional functions of assessment. In many systems, including Vietnam, exams serve not only pedagogical purposes but also certification, selection, and social mobility. Their deep entrenchment within educational structures makes them resistant to change, even when their limitations are widely acknowledged.

In response, a growing body of scholarship in TESOL and educational assessment has questioned the dominance of summative testing, advocating for formative and sustainable approaches to assessment that extend beyond immediate performance (Boud & Falchikov, 2007; Carless, 2015). More recent work highlights learner engagement in assessment, particularly through co-regulation of learning and learners' capacity to make evaluative judgements (Andrade & Brookhart, 2019; Tai et al., 2018). Within this broader shift, "no-exam" classrooms have emerged as a provocative alternative.

However, such approaches are often framed in overly optimistic terms. Rather than treating them as solutions, this article argues for a reconceptualization of assessment itself, one that seeks to manage its enduring tensions rather than to resolve them. By critically examining the shift from exam-based to evidence-based assessment, it explores both the potential and the limitations of no-exam classrooms in the context of the EFL-to-ESL transition.

2. Rethinking Assessment: From Exams to Evidence

The notion of "no-exam" classrooms is frequently interpreted as a rejection of assessment. In practice, however, it represents a transformation in how assessment is conceptualized and enacted. Rather than eliminating evaluation, these approaches redistribute it across time, tasks, and interactions, shifting from a single, high-stakes event to an ongoing process of evidence collection and interpretation. This shift can be conceptualized as a movement from a "snapshot" model of assessment to a longitudinal, evidence-based model, as summarized in Table 1. The two outer columns are presented as contrasting ideal types rather than mutually exclusive options; the third column sketches a hybrid model that draws selectively on both and anticipates the integrated framework developed in Section 4.

Table 1.

From Snapshot Testing to Longitudinal Evidence

Traditional Exam Model	Evidence-Based (No-Exam) Model	Hybrid (Integrated) Model
Single high-stakes performance	Continuous, low-stakes evidence	Periodic anchor points set within continuous evidence
Time-limited, decontextualized tasks	Contextualized, interaction-based tasks	Authentic tasks under partly standardized conditions
Focus on accuracy and correctness	Focus on communicative effectiveness	Accuracy and communicative effectiveness weighted together
Product-oriented	Process + product oriented	Process and product, with selective benchmarking
One-time measurement	Development over time	Development tracked through recurring checkpoints

Traditional exam-based assessment operates within a measurement paradigm, assuming that language ability can be quantified through standardized tasks administered under controlled conditions. This approach prioritizes reliability and comparability, ensuring that all

learners are assessed using the same criteria. However, such standardization often requires the simplification of complex constructs, reducing communicative competence to discrete, measurable components (Fulcher, 2010).

This reduction has important consequences. By focusing on what can be easily measured, exams may neglect aspects of language use that are more difficult to quantify, such as interactional competence, pragmatic appropriateness, and strategic communication. As a result, there is a risk of construct under-representation, where the assessment fails to capture the full range of abilities it claims to measure (Messick, 1989).

In contrast, evidence-based assessment models expand the range of evidence used to evaluate learning, incorporating spoken interactions, written outputs, reflective accounts, peer feedback, and teacher observations. This allows for a more comprehensive and context-sensitive representation of language ability. However, this shift from measurement to evidence introduces new challenges. Unlike exam scores, which produce clear and comparable outcomes, evidence-based assessment relies on interpretive judgment, raising questions about reliability, transparency, and fairness.

Furthermore, this shift redefines assessment as an interpretive and context-dependent practice rather than a purely technical process of measurement. In traditional systems, knowledge is treated as something that can be sampled and generalized across contexts, aligning with a positivist view of assessment. However, when assessment is based on multiple forms of evidence, this epistemological foundation becomes less stable.

One key issue is the problem of commensurability. Different forms of evidence, such as interaction, written production, and reflection, do not easily translate into a single metric. Unlike test scores, which are designed to be comparable, these forms of evidence are inherently heterogeneous. This raises the question of whether assessment should prioritize comparability or embrace the richness and contextuality of evidence, even at the expense of standardization.

This tension is particularly salient in ESL-oriented contexts, where language use is situated and context-dependent. If communicative competence varies across contexts, interlocutors, and purposes, then reducing it to a single score may be inherently reductive. Evidence-based approaches respond to this by embracing variability rather than controlling for it, but this also complicates decision-making in institutional settings where comparable outcomes are required.

Another critical issue concerns the temporal dimension of assessment. Traditional exams capture performance at a specific moment, whereas evidence-based assessment tracks development over time. While this allows for a more nuanced understanding of learning, it raises questions about what counts as representative performance, that is, whether assessment should privilege peak performance, typical performance, or improvement.

Finally, the redistribution of assessment across time and contexts alters the relationship between assessment and learning. In exam-driven systems, assessment often functions as a terminal event. In evidence-based models, it becomes embedded within ongoing activity. While this integration is pedagogically desirable, it also blurs the boundary between formative and summative assessment, raising questions about how and when evaluative judgments should be made.

3. Theoretical and Critical Perspectives

The move toward no-exam assessment is often justified through established theoretical frameworks, including sociocultural theory, constructivism, and Assessment for Learning.

While these perspectives provide strong support for alternative approaches, they also reveal underlying tensions that complicate their implementation.

From a sociocultural perspective, learning is inherently social and mediated through interaction (Vygotsky, 1978). Language competence emerges through participation in communicative practices, where meaning is co-constructed between individuals. This challenges the validity of traditional exams, which isolate learners from interactional contexts. However, it also complicates the notion of individual assessment. If performance is co-constructed, it becomes difficult to determine the extent to which it reflects individual ability. This raises important questions about attribution, fairness, and the role of assessment in evaluating individual learners.

Constructivist theory further conceptualizes learning as an active process of meaning-making (Piaget, 1970; Boud & Falchikov, 2007). Knowledge is not reproduced but constructed through engagement with tasks and experiences. This challenges the epistemological assumptions of exam-based assessment, which often treats knowledge as stable and transferable. However, constructivism also implies that learning is inherently variable, shaped by individual experiences and interpretations. This variability complicates the application of standardized assessment criteria and raises questions about how consistency can be achieved without oversimplifying the learning process.

Assessment for Learning (AfL) emphasizes the central role of feedback in supporting learning (Black & Wiliam, 1998). Carless (2015) extends this perspective by highlighting dialogic feedback, where learners actively engage with feedback processes. More recent work emphasizes learners' capacity to interpret and use feedback effectively, particularly through the development of evaluative judgement (Tai et al., 2018). These perspectives support the integration of assessment into learning processes and align with the principles of no-exam classrooms.

However, they also introduce new challenges. The emphasis on learner agency assumes that learners possess the skills and dispositions necessary to engage with feedback. In contexts where students are accustomed to exam-driven systems, this assumption may not hold. Without explicit support, learners may struggle to interpret feedback or to translate it into meaningful action. As a result, the effectiveness of no-exam approaches may depend on the development of new forms of assessment literacy among both teachers and students.

From a validity perspective, Messick (1989) argues that assessment must accurately represent the construct it claims to measure. In communicative language teaching, this construct is inherently complex and context-dependent. While no-exam approaches attempt to enhance construct validity by incorporating multiple forms of evidence, they simultaneously reduce standardization. This creates a tension between validity and reliability that cannot be fully resolved.

Moreover, the shift toward evidence-based assessment reconfigures power relations within the classroom. Traditional exams position teachers as evaluators and students as subjects of evaluation. In contrast, no-exam approaches emphasize collaboration, reflection, and shared responsibility. While this may be pedagogically desirable, it also challenges established roles and may create uncertainty for both teachers and learners.

Taken together, these theoretical perspectives suggest that no-exam assessment is not a straightforward solution, but a reconfiguration of existing dilemmas. It foregrounds issues of interpretation, agency, and context, while introducing new challenges related to consistency, scalability, and institutional alignment.

A further layer of complexity emerges when these theoretical perspectives are considered not only as justifications for alternative assessment, but as lenses for critique. While sociocultural theory emphasizes interaction and mediation, it does not necessarily provide clear guidance on how to evaluate individual performance within collaborative contexts. If learning is distributed across participants and tools, then assessment must account for this distribution. However, most educational systems continue to require individual certification, creating a structural tension between theory and practice.

The same critique applies to constructivist accounts. Because they treat learning trajectories as individualized, they support personalized, process-oriented assessment but sit uneasily with the notion of common standards. If learning is inherently individualized, then the application of uniform assessment criteria may risk misrepresenting learners' achievements. This raises a critical question: to what extent should assessment prioritize comparability over representational accuracy?

The concept of feedback literacy further complicates this picture. While Tai et al. (2018) emphasize the development of evaluative judgement, which underpins learners' ability to interpret and use feedback effectively, this capacity cannot be assumed and must be explicitly taught and scaffolded. In contexts where students have been socialized into exam-oriented systems, the transition to feedback-rich environments may require significant adjustment. Without such support, learners may struggle to engage with feedback, leading to superficial or ineffective use of assessment information.

Moreover, the emphasis on learner agency in no-exam approaches may obscure the continued influence of institutional and cultural structures. While students may be encouraged to take ownership of their learning, their actions remain shaped by broader expectations, including grading systems, curriculum requirements, and future opportunities. In this sense, agency is not simply a matter of individual choice, but is mediated by systemic constraints.

From a critical perspective, it is also important to consider how power operates within assessment systems. Traditional exams centralize authority in the hands of institutions and examiners, while no-exam approaches redistribute some of this authority to teachers and learners. However, this redistribution is not neutral. It may create new forms of inequality, particularly if learners differ in their ability to engage with reflective and self-directed processes. Students with stronger metacognitive skills or greater access to support may benefit disproportionately, raising concerns about equity.

Finally, the shift toward evidence-based assessment can be understood as part of a broader movement toward what Boud and Falchikov (2007) describe as sustainable assessment. This approach emphasizes the development of learners' capacity to evaluate their own work beyond formal education. While this represents an important goal, it also requires a redefinition of success in education, with a shift from performance on tests to the ability to engage in ongoing learning. This has far-reaching implications not only for assessment practices, but for the purposes of education itself.

4. Toward a Reconceptualised Assessment Framework

Rather than framing the debate as a binary choice between exam-based and no-exam approaches, a more productive perspective is to reconceptualize assessment as a system that integrates multiple forms of evidence while remaining responsive to institutional and pedagogical demands. The question is therefore not whether exams should be retained or removed, but how different assessment practices can be combined to better reflect the complexity of language learning in ESL-oriented contexts.

This reconceptualization requires a shift from viewing assessment as a tool for measurement to understanding it as a system of practices that mediate learning, evaluation, and decision-making. In traditional exam-based models, assessment is largely separated from instruction and designed to produce standardized outcomes. In contrast, evidence-based approaches embed assessment within learning processes, emphasizing development, interaction, and feedback. However, as the previous discussion has shown, neither approach is without limitations.

At the core of this issue lies a set of persistent tensions that cannot be fully resolved, but must instead be managed. One such tension lies between validity and reliability. While evidence-based assessment enhances construct validity by capturing authentic and context-sensitive language use, it often reduces standardization, making it more difficult to ensure consistency across learners (Messick, 1989). Conversely, traditional exams achieve high reliability through controlled conditions, but risk under-representing the construct of communicative competence.

A related tension exists between authenticity and comparability. Tasks that closely resemble real-world language use are inherently variable, shaped by context, interlocutors, and purpose. While this variability reflects the nature of language, it complicates efforts to compare learners on a common scale. Similarly, there is a tension between flexibility and accountability. Evidence-based systems allow for adaptation to learners' needs and trajectories, but institutions often require clear, comparable outcomes for purposes of certification and progression.

Rather than attempting to eliminate these tensions, a reconceptualised approach to assessment makes them explicit and seeks to balance them through principled design. To address these tensions, a reconceptualised framework for ESL assessment can be proposed, integrating key principles that balance validity, reliability, and pedagogical relevance (see Table 2).

Table 2.

A Reconceptualised Framework for ESL Assessment: Key Design Principles and Implications

Principle	Description	Implication
Multimodal Evidence	Combines product, process, and interaction	Reduces reliance on single performance
Longitudinal Assessment	Tracks development over time	Reflects learning as progression
Analytic Criteria	Uses clear rubrics	Enhances transparency and consistency
Feedback Integration	Embeds feedback in learning cycles	Supports continuous development
Learner Agency	Involves reflection and self-assessment	Promotes self-regulated learning

Table 2 outlines a reconceptualised assessment framework that seeks to balance competing demands within ESL-oriented contexts. Rather than privileging a single form of evidence, the framework emphasizes multimodality, drawing on product, process, and interactional data to provide a more comprehensive representation of language ability. This approach aligns with calls to enhance construct validity by capturing the complexity of communicative competence.

At the same time, the inclusion of analytic criteria and structured rubrics reflects the need for transparency and consistency in evaluation. While evidence-based approaches often prioritize authenticity, without clear criteria they risk becoming opaque or inconsistent.

Analytic frameworks therefore play a crucial role in mediating between flexibility and standardization, providing a shared basis for interpretation while allowing for contextual variation.

The integration of feedback as a central component of assessment further shifts the focus from judgment to development. Rather than functioning solely as an evaluative endpoint, assessment becomes part of an ongoing cycle of feedback and improvement (Carless, 2015). However, as noted earlier, this shift depends on learners' capacity to engage with feedback effectively. The emphasis on learner agency must therefore be accompanied by explicit support for the development of evaluative judgement (Tai et al., 2018).

Importantly, the framework does not resolve the tensions identified earlier, but redistributes them. For example, increasing the use of authentic, interaction-based tasks may enhance validity but reduce comparability across learners. Similarly, promoting learner autonomy may support engagement, but also introduce variability in how assessment is enacted. These trade-offs are not shortcomings of the framework, but inherent features of assessment in complex learning environments.

In practice, this suggests the need for hybrid models that combine elements of both traditional and alternative approaches. Rather than eliminating exams entirely, they can be repositioned within a broader system of evidence, where their role is complemented by ongoing, process-oriented assessment. Such models allow for the retention of standardization where necessary, while also incorporating richer forms of evidence that better reflect communicative competence.

A brief scenario illustrates how such a model might convert heterogeneous evidence into a defensible final grade. Suppose an institution requires a single end-of-term mark for an ESL speaking course. A teacher might allocate that mark across three weighted strands: 40% to interactional evidence gathered over the term (recorded pair and group tasks, judged against an analytic rubric for fluency, interactional management, and pragmatic appropriateness), 30% to a portfolio of process evidence (drafts, self-assessments, and responses to feedback), and 30% to a standardized end-of-term task that satisfies the institution's demand for comparability. Because the interactional strand is the least standardized, it is the one most in need of safeguards: at least two interaction samples are rated, a second rater reviews a sample of borderline cases, and the rubric is shared with learners in advance so that judgements remain transparent. Where the standardized task and the interactional record diverge sharply, the discrepancy is not averaged away but flagged for review, prompting the teacher to ask whether the exam under-represented the learner's communicative competence or whether the continuous evidence was inflated by uneven task conditions. The final grade is therefore not a precise calculation but a principled, auditable judgement, with the weighting scheme making explicit how interactional data and institutional metrics have been balanced.

Ultimately, a reconceptualised approach to assessment requires a shift in how assessment is understood: not as a problem to be solved through technical refinement, but as a practice that must be continually negotiated in relation to changing educational goals and contexts. In ESL-oriented environments, where language use is inherently dynamic and situated, this negotiation becomes central to the design of meaningful and credible assessment systems.

5. Conclusion: Reframing Assessment in the EFL-to-ESL Transition

The transition from EFL to ESL contexts represents more than a shift in language policy; it entails a fundamental rethinking of what it means to know and use a language. As

communicative competence becomes the central goal of language education, assessment practices must evolve to reflect this shift. Traditional exam-based systems, while efficient and standardized, may no longer be sufficient to capture the complexity of language use in real-world contexts.

No-exam classrooms offer a compelling response to this challenge by emphasizing continuous, evidence-based assessment aligned with learning processes. However, as this paper has argued, they should not be viewed as a complete replacement for exams, but as part of a broader reconceptualization of assessment. They highlight important limitations of traditional systems, but also introduce new tensions related to reliability, workload, learner readiness, and institutional constraints.

In the context of the EFL-to-ESL transition, the key challenge is not whether to retain or remove exams, but how to design assessment systems that are both pedagogically meaningful and institutionally viable. This requires a shift from viewing assessment as a tool for measurement to understanding it as a complex, context-dependent practice that supports learning while maintaining credibility.

Ultimately, the future of TESOL assessment lies not in choosing between exams and no-exam approaches, but in developing integrated systems that balance competing demands. By acknowledging and engaging with these tensions, rather than attempting to eliminate them, language educators can move toward more valid, equitable, and sustainable forms of assessment in ESL-oriented contexts.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74.

Andrade, H. L., & Brookhart, S. M. (2019). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 1–23. https://doi.org/10.1080/0969594x.2019.1571992

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. Routledge.

Carless, D. (2015). *Excellence in university assessment*. Routledge.

Fulcher, G. (2010). *Practical language testing*. Routledge.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Macmillan.

Piaget, J. (1970). *Science of education and the psychology of the child*. Orion Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Bionote

Ms. Nguyen Thi Thanh Thao is a full-time lecturer at the International Institute, Ho Chi Minh City University of Economics and Finance, Vietnam. She is currently a PhD student in English Language Studies at Suranaree University of Technology, Thailand. Her research interests include teaching methodologies, language assessment, and technology-enhanced language learning.

Ms. Le Thi Dieu Duyen is a full-time lecturer at the Faculty of Foreign Languages & Oriental Studies, Yersin University of Da Lat, Vietnam. She is currently a PhD student in English Language Studies at Suranaree University of Technology, Thailand. Her research interests include discourse analysis, sociolinguistics, and technology-enhanced language learning.

Ms. Chau Thi Hong Tham is a lecturer at Nguyen Tat Thanh University, Vietnam, specializing in English language teaching methodologies. She is currently a PhD student in English Language Studies at Suranaree University of Technology, Thailand. Her research interests focus on teacher professional development, teacher identity construction, and language ideology in educational contexts.

Mr. Bui Thanh Tinh is an EFL lecturer who has been working at Dong Thap University for more than eight years. He is currently a PhD student in English Language Studies at Suranaree University of Technology, Thailand. His research interests focus on innovative approaches to English language teaching, and language testing and assessment.

Mr. Tran Van Doan is a full-time lecturer at the Faculty of Foreign Languages, Ho Chi Minh City University of Industry and Trade, Vietnam. He is currently a PhD student in English Language Studies at Suranaree University of Technology, Thailand. His research interests include language mindset theory, language assessment, and technology-enhanced language learning.

#142 ENGLISH-MAJOR STUDENTS' READINESS FOR REAL WORLD ENGLISH USE IN AN ESL-ORIENTED CONTEXT: A CASE STUDY AT HUFLIT

Loan Nguyen Thi Thanh[1]

Abstract

This study aims to investigate English-major students' readiness for real-world English use in an ESL-oriented context at HUFLIT University, focusing on three dimensions: communicative competence, willingness to communicate, and L2 self-confidence. Data were collected from 160 English-major students using a questionnaire. Descriptive statistics were employed to examine learners' overall readiness and their distribution across the three dimensions. In addition, one-way ANOVA was conducted to explore differences in readiness across academic year and self-rated English proficiency, while Pearson correlation analysis was used to examine the relationships among the three dimensions. The findings revealed that students reported a high level of perceived readiness toward ESL orientation. While no significant differences were found across academic years, the level of readiness varied significantly according to English proficiency levels. Furthermore, strong positive correlations were identified among the three dimensions. These findings suggest that enhancing learners' readiness to use English as a second language requires not only the development of language competence but also the fostering of confidence and a supportive environment that encourages active communication. The study examines perceived readiness rather than actual real-world English performance.

***Keywords*:** learner readiness, ESL-oriented context, willingness to communicate, L2 self-confidence

1. Introduction

English has increasingly become a global means of communication, particularly in academic, professional, and international contexts. In response, language education has gradually shifted from a focus on grammatical knowledge toward the ability to use English for real-world communication. In Vietnam, English has traditionally been taught as a foreign language within the Expanding Circle (Kachru, 1985). However, recent policy developments indicate a shift toward a context in which English is expected to function as a second language across a wide range of educational system (Decision No. 2371/QĐ-TTg, 2025). Despite this orientation, English language education in Vietnam remains largely exam-oriented, with limited opportunities for authentic communication (Nguyen et al., 2018). As a result, learners may not be sufficiently prepared to use English in real-life contexts.

The concern of using English as second language is especially relevant in higher education, where English-major students are expected to become future users. English language instruction in Vietnamese universities is typically divided into English-major and non-English-major programs, with the former placing English as the central focus of study (Hoang, 2010). In order to apply English in various working areas including translation, teaching, and business, English major students are educated not only language skills but also knowledge in linguistic, culture and literature. However, despite this intensive training,

[1]*Nguyen, Thi Thanh Loan (Corresponding author), Lecturer, Department of English Language Studies, Faculty of Foreign Languages, HUFLIT, Ho Chi Minh City, Vietnam; Email: loanntt@huflit.edu.vn*

persistent limitations remain. Studies indicate that only a small proportion of graduates achieve sufficient proficiency for professional use, while others continue to face difficulties in speaking, reading, and writing (Vo et al., 2018; Tran & Nguyen, 2023; Dang et al., 2021).

While previous studies have largely focused on language proficiency and learning outcomes, limited attention has been paid to learners' readiness to use English in real-world contexts. In other words, existing studies in this field have mainly examined what English-major students can do linguistically, while this study examines whether they feel ready to use English beyond the classroom in an ESL-oriented context. Moreover, little is known about whether English-major students, who are expected to function as future users of English, are adequately prepared for this shift. Therefore, this study aims to investigate English-major students' readiness for real-world English use in an ESL-oriented context.

The current study draws on three research questions:

1. To what extent are English-major students ready for real-world English use in an ESL-oriented context?

2. To what extent does learners' readiness differ across academic year, and self-rated English proficiency?

3. What are the relationships between communicative competence, willingness to communicate, and L2 self-confidence in shaping learners' readiness?

2. Literature Review

2.1. From EFL to ESL in Language Education

Global Context: Early efforts to classify varieties of English in Second Language Acquisition (SLA) research are often associated with Kachru's (1985, 1992) Three Circles model, which categorizes English use into the Inner, Outer, and Expanding Circles, broadly aligning with English as a Native Language (ENL), English as a Second Language (ESL), and English as a Foreign Language (EFL). According to Schneider (2003, 2007), in Outer Circle countries where English functions as a second language, it has been historically institutionalized and is used across various domains of daily life. In contrast, in Expanding Circle contexts, English is typically learned as a foreign language and is mainly confined to formal education, although it may also serve as a lingua franca in international communication. Furthermore, in SLA, a foreign language is introduced as "the learning of a nonnative language in the environment of one's native language", while a second language refers to "the learning of nonnative language in the environment in which that language is spoken" (Gass & Selinker, 2001, p. 5).

However, the distinction between EFL and ESL is fluid and context dependent. A study of International Corpus of Learner English (Gilquin & Granger, 2011) reveals that learners, who are traditionally categorized as EFL users, sometimes experience learning conditions closer to ESL contexts. The blurring is contributed by several factors including academic years, teaching methods, EMI use, teacher input quality, the exposure to English outside the classroom, and English language media access. Therefore, EFL and ESL are better understood as points along a continuum rather than as clear separations.

Local Context: Traditionally, Vietnam has been categorized in the Expanding Circle, where English has been taught in school as formal subjects (Kachru, 1985). After the 1986 Economic Reforms, the country increased integration into the global economy, considering English as a tool for economic development and international engagement (Do, 2000). Later,

the National Foreign Language Project 2020 has further promoted English instruction through standardized proficiency benchmarks, and alignment with international frameworks (Nguyen et al., 2018). More recently, a noticeable shift in policy orientation has emerged, the Decision No. 2371/QĐ-TTg, formally introduces ESL within the educational system. According to this direction, English is expected to function as a second language, supporting communication in various domains. This change should be viewed as part of a continuum-based development. However, a key question regarding learners' readiness toward using English beyond the classroom still has been underdetermined.

2.2. English-major Students as Future English Users

In Vietnamese higher education, English language studies are commonly divided into English- major and non-English major programs. While non-English major program is designed for students in different majors using English as a support, English- major students undertake English as the core component of their studies (Hoang, 2010). This program emphasizes the practical application of English in various professional contexts such as translation, education, administration, media, tourism, and international business. Therefore, these students are expected to achieve strong communicative competence across a range of situations, including cross-cultural and international workplaces. During four academic years, they are trained to integrate language skills with many soft skills in respond to the demand of globalization. In fact, they are considered as future English users who English as a tool for international communication and collaboration.

However, only a small number of English- major graduates reach a level sufficient for professional use. Particularly, students find difficulties with speaking because of limited exposure to authentic communicative environments (Vo et al., 2018). They also have struggle with reading comprehension due to the lack of vocabulary, background knowledge, and motivation (Tran & Nguyen, 2023). In addition, current studies indicate such English- major students' challenges in linguistic competence in their overall communicative ability (Dang et al., 2020). These limitations can be explained by the remains of exam-oriented, grammatical accuracy and test performance emphasis (Nguyen et al, 2018). The lack of real-life English environments further constrains their ability to use the language confidently and effectively as a second language. These current conditions raise a controversial question: While there is an increasing effort on ESL-oriented use of English in Vietnamese educational system, whether English- major students are fully prepared for the demands of ESL-oriented context?

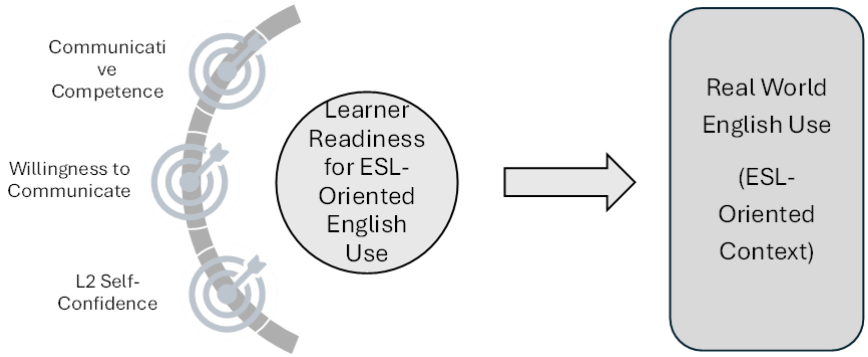
2.3. Learner Readiness for Real-world English Use

Learner readiness is commonly defined as learners' ability to acquire knowledge and initiate behavioral changes that lead to effective learning outcomes (Chorrojprasert, 2020). Conceptually, learner readiness is a multidimensional construct encompassing both psychological–educational and developmental dimensions (Lightbown & Spada, 1999). The former involves learners' cognitive abilities, motivation, attitudes, and self-regulation, while the latter refers to their stage-based capacity to process and acquire linguistic features in sequence.

Building on this perspective, this study adopts a conceptual framework in which learner readiness toward an ESL-oriented paradigm is conceptualized as the integration of communicative competence, willingness to communicate, and self-confidence use of English in real world communication.

Figure 1

Conceptual framework of learner readiness for ESL-oriented English use.



The three dimensions are not independent constructs but dynamically interact to shape learners' readiness for real-world English use. Communicative competence provides the linguistic and sociocultural resources necessary for effective communication. However, possessing competence alone does not guarantee actual language use. Learners must also perceive themselves as capable of communicating successfully. In this sense, communicative competence can foster L2 self-confidence because learners who believe they have adequate language ability are more likely to feel secure in engaging in interaction. Higher levels of self-confidence, in turn, tend to increase learners' willingness to communicate, as confident learners experience less anxiety and are more prepared to initiate and sustain conversations. Willingness to communicate then serves as the immediate antecedent of actual language use by activating learners' linguistic resources in authentic situations. Therefore, learner readiness in an ESL-oriented context should be understood as the result of a dynamic interplay among competence, confidence, and willingness to communicate rather than the product of any single dimension.

Communicative competence refers to the ability to use English in real life situations successfully. According to Savignon (2018), language proficiency should be evaluated based on learners' ability to communicate in meaningful contexts rather than grammatical or lexical tests. This ability varies depending on communicative purposes and practical contexts. Therefore, it requires learners to initiate interaction and respond appropriately in authentic situation (McGroarty, 1984). Although communicative competence includes sociocultural norms and patterns of interaction within the target language community, without supporting factors, having such ability does not ensure that learners will use language effectively (Clément, 1980).

The second dimension of learner readiness in an ESL-oriented context is willingness to communicate, which is considered as one of the most predictors of actual L2 use. Learners who are willing to communicate tend to join in interaction and practice the target language actively (MacIntyre et al., 1998). Therefore, this factor has been positioned as a primary goal of ESL education because the aim of this orientation is to encourage learners to use English for meaningful situations. In this case, willingness to communicate is a bridge between communicative competence and performance. Although communicative competence encompasses linguistics, learners must have willingness to communicate to activate these competencies in real interaction (Scarcella et al., 1990).

Although willingness to communicate is significant, it strongly depends on learners' level of L2 self-confidence, which can be understood as learners' perception of their own competence and ability to successfully use the target language (Tunçel, 2015). Self-confidence

reflects both self-assurance and cognitive evaluation of one's communicative capability. Learners who perceive themselves as good communicative competence tend to engage in interaction because self-confidence has been identified as a key predictor of willingness to communicate (Bensalem & Thompson, 2021). In ESL-oriented contexts, self-confidence becomes more crucial because without such confidence, learners' ability to translate linguistic knowledge into meaningful communication may be constrained.

3. Methodology

To capture the general pattern of learner readiness, this study adopted a quantitative research design. A quantitative approach was considered appropriate because the study aimed to measure students' perceived readiness across three predefined dimensions and to examine differences and relationships among variables in a systematic and objective manner. The use of a survey enabled the collection of standardized and quantifiable data from a relatively large number of participants, thereby facilitating statistical comparisons across groups and identifying relationships among communicative competence, willingness to communicate, and L2 self-confidence.

Participants

The participants of this study were 160 English-major students at HUFLIT, regardless of gender, academic year, or self-rated English proficiency. A convenience sampling approach was employed because it enabled access to available participants within the institutional context and facilitated data collection from a relatively large number of students. However, because participation depended on accessibility and willingness to participate, the sample cannot be considered fully representative of all English-major students at HUFLIT or of Vietnamese English-major students in general. Therefore, the findings should be interpreted within the specific context of the present study.

Instrument

This study used a bilingual survey presented in both English and Vietnamese to ensure clarity and facilitate participants' comprehension. There were two main sections, in which section two included three sub dimensions. Initially, there were 21 items in the survey. However, after piloting and cleaning, the final number of included items was 17.

Table 1.

Structure of Survey After Pilot

Survey (17 items in total)

Demographic Information (3 items)			Learner Readiness (14 items)		
Gender (1 item)	Academic Year (1 item)	Self-rated English Proficiency (1 item)	Communicative Competence (4 items)	Willingness to Communicate (5 items)	L2 Self-Confidence (5 items)
			Cronbach's alpha = .732	Cronbach's alpha = .924	Cronbach's alpha =.895

The first sections included 3 items. These items were used to collect demographic information, including gender, academic year, and self-rated English proficiency. Reponses from these items provided background information about the participants and supporting the description of the sample.

The second section of the questionnaire focused on measuring learners' readiness for real-world English use. This section employed a 5-point Likert scale ranging from strongly disagreement to strongly agree. The survey items were developed based on the study's conceptual framework, and were organized into three dimensions: Communicative competence, willingness to communicate, and L2 self-confidence. Initially, this section consisted of 18 items, with 6 items for each dimension. Then, a pilot study was conducted to evaluate the clarity and comprehensibility of the items, as well as the reliability of the scale. Based on the pilot results, 3 items were removed to ensure that the Cronbach's alpha for each dimension exceeded 0.70 and that no corrected item–total correlation fell below 0.30.

Data Analysis Process

Firstly, the data collected were analyzed using IBM SPSS Statistics version 27. First, the internal consistency of the survey was re-examined using Cronbach's alpha, with all scales demonstrating acceptable reliability (Cronbach's alpha >.7).

Secondly, descriptive statistics, including mean scores and standard deviations, were calculated to determine the overall level of learners' readiness for real-world English use. The analysis was then extended to compare mean scores across the three dimensions to identify the most and least developed aspects of readiness.

Thirdly, a one-way ANOVA was conducted to examine whether there were statistically significant differences in learners' readiness across different groups including academic years and English self- rated proficiency level.

Finally, correlation analysis was employed to explore the relationships among communicative competence, willingness to communicate, and L2 self-confidence of learners' readiness.

4. Results

Reliability Check

The internal consistency of the three dimensions was examined using Cronbach's alpha. The results showed that all constructs demonstrated high levels of reliability, with alpha coefficients exceeding the recommended threshold of 0.70, ranging from .833 to .906. In addition, all items across the three dimensions reported Corrected Item–total Correlations above 0.30, indicating that each item contributed meaningfully to its respective construct. Overall, these findings confirm that the instrument is reliable and appropriate for further statistical analysis.

Descriptive Analysis

The overall mean score of students' readiness for real-world English use was 3.73 (SD = 0.70), indicating a high level of readiness. However, the standard deviation suggests a certain degree of variation, which showed that not all learners demonstrate the same readiness level.

Table 2

General Learner Readiness for Real-World English Use

Dimensions of Learner Readiness	N	Min.	Max.	Mean	SD
Communicative Competence	160	2.00	5.00	3.7188	.68322
Willingness to Communicate	160	1.6	5.00	3.8925	.75565
L2 Self-Confidence	160	1.4	5.00	3.5662	.85773

Overall	160	2.17	5.00	3.7258	.70058

As shown in Table 2, learner readiness varies across three dimensions. While willingness to communicate has the highest mean score (M = 3.89, SD = .76) and followed by communicative competence dimension (M=3.72, SD=.683), L2-self-confidence positions at the bottom of the rank (M = 3.57, SD = .86). This pattern suggests that although learners are generally willing to engage in communication, they still lack confidence of using English in real world context. The result highlights the role of psychological factors in shaping English using behavior.

One-way ANOVA Analysis

Regarding academic year, Levene's test confirmed that the assumption of homogeneity of variances was met (p = .817). The one-way ANOVA showed no statistically significant difference in overall readiness among students from different academic years (F (3, 156) = 1.54, p = .206).

Post-hoc analysis using Tukey's HSD test also revealed no significant differences between any pairs of groups (all p > .05), suggesting that students' readiness remains relatively consistent across academic years.

Table 3

One-way ANOVA by English Proficiency Level

Proficiency Level	N	Mean	SD
Beginner	6	3.6667	1.044
Intermediate	88	3.5617	.6841
Upper-intermediate	61	3.8973	.6123
Advanced	5	4.5933	.5648
Total	160	3.7258	.7006
One-way Anova		F=5.895	p<.001

In contrast, a significant difference was found across levels of self-rated English proficiency. Levene's test indicated that the assumption of homogeneity of variances was satisfied (p = .094). The ANOVA results revealed a statistically significant difference among the groups (F (3, 156) = 5.90, p < .001), indicating that learners' readiness varies according to their English proficiency levels.

Post-hoc analysis using Tukey's HSD test showed that students at the intermediate level reported significantly lower readiness compared to those at the upper-intermediate (p = .016) and advanced levels (p = .006), while no significant differences were found among the remaining groups.

Correlation Analysis

Table 4

Pearson Correlations among Learner Readiness Dimensions

	Communicative Competence	Willingness to Communicate	L2 Self-Confidence
Pearson Correlation	1	.721**	.781**

Communicative Competence	Sig. (2-tailed)		.000	.000
Willingness to Communicate	Pearson Correlation	.721 **	1	.756**
	Sig. (2-tailed)	.000		.000
L2 Self-Confidence	Pearson Correlation	.781**	.756**	1
	Sig. (2-tailed)	.000	.000	

***. Correlation is significant at the 0.01 level (2-tailed).*

Pearson correlation analysis was conducted to examine the relationships among the three dimensions of learner readiness. The results revealed significant positive correlations among all variables ($p < .001$). Specifically, communicative competence was strongly correlated with willingness to communicate ($r = .721$, $p < .001$) and self-confidence ($r = .781$, $p < .001$). In addition, willingness to communicate was also strongly associated with self-confidence ($r = .756$, $p < .001$).

The three dimensions are interconnected closely and the relationship among them is positive, which demonstrates the study's conceptual framework. Based on this finding, learner readiness is a multidimensional construct where communicative competence, willingness to communicate, and L2 self-confidence influence each other.

5. Discussion

RQ1

English-major students' readiness for real-world English use

The findings suggest that English-major students at HUFLIT generally perceived themselves as being ready for real-world English use. The relatively high level of perceived readiness may reflect students' positive learning experiences and their exposure to communicative activities during their studies. Therefore, these explanations should be considered tentative rather than definitive conclusions. As Chorrojprasert (2020) explains in a similar study conducted in Thailand, rather than staying focuses on language knowledge, creating a real communicative environment and motivate learners to take part in is the core idea to push their readiness in practicing language.

The comparatively lower level of L2 self-confidence deserves particular attention. Although students reported a high willingness to communicate, limited confidence may prevent them from translating this willingness into actual language use. Learners may be psychologically prepared to engage in communication but still hesitate to participate in authentic interactions because they doubt their linguistic abilities. This finding suggests that willingness to communicate alone is insufficient to ensure real-world English use if learners do not perceive themselves as capable language users.

However, the variation among learner readiness indicates a uniform distribution. While large numbers of learners have well-prepared in English using, some of them still face challenges when shifting from English used in classroom, EFL, to a more authentic English beyond the classroom, ESL-oriented environment. Therefore, readiness for English real-world communication is a complex construct that combines not only linguistic competence but also psychological and contextual factors.

RQ2

Learner readiness differs across academic years and English proficiency

The level of leaner readiness toward using English in real-world context does not develop effectively in a linear manner because the findings show there is no significant differences within learners' year of study. It demonstrates that even senior students might have less readiness than junior students if they remain a limitation of authentic communication opportunities. In other words, simply advancing through academic years within a training program does not automatically translate into greater preparedness for using English as a second language.

In contrast, there is a great difference across self-rated English proficiency levels which is noted as one of a critical factor in shaping learner readiness. Students, who determine themself at the beginner and intermediate level, report lower readiness compared to those, who refer themselves as upper-intermediate and advanced English user. This is because perception can shape decision making (Delgado, 2021). When learners believe that they have good control over their language, they are more likely to feel confident and engage in communication. In contrast, learners with lower proficiency evaluation may experience hesitation that might hinder their willingness to communicate. Rather than indicate the practical English proficiency of learners, the finding focuses on the way they self-evaluate their L2 capacity. Their real current English level might not as the same as they are self-rated, but it reflects their belief about their ability leading to their willingness to communicate. This finding aligns with MacIntyre's (1994) Willingness to Communicate model, referring that when learners perceive their communicative competence, they tend to decrease the communication anxiety and increase communicative motivation, resulting in ready to raise their voices.

Overall, the significant factor to improve learner readiness here is the combination of language competence and learners' awareness of their ability, rather than relying on academic progression.

RQ3

The relationships between communicative competence, willingness to communicate, and L2 self-confidence in shaping learners' readiness

Learner readiness should not be considered as a single construct because there is a strong positive relationship among three dimensions. Particularly, communicative competence is a central dimension that leads to confident feeling and engagement in communication. For instance, when learners aware of their English competence, they belief that they can use English effectively. This motivation plays as a momentum that push learners to participant in real English communication context. The more they are willing to engage in such situations, the more readiness they are building. This interconnection supports the assumption that linguistic ability and perception set a foundation upon other behavioral factors are built.

The strong association between self-confidence and willingness to communicate further highlights the importance of psychological factors in SLA. Even when learners have well-prepared linguistic knowledge, their confident feelings might heavily affect their communicative engagement. Based on the findings of this study, self-confidence acts like a facilitator that encourages learners to activate their language abilities into practical communicative behaviors. Therefore, enhancing learner readiness toward using English as second language requires not only language competence improvement but also supportive environment where learners can foster their confidence.

Furthermore, it is important to note that proficiency in the present study was self-rated rather than objectively measured. Therefore, the significant differences across proficiency groups may reflect students' perceptions of their own language ability rather than their actual

level of English proficiency. This interpretation strengthens the theoretical argument of the study because perceived competence is closely associated with L2 self-confidence and willingness to communicate. Learners who believe they are competent are more likely to feel confident and engage in communication, regardless of their objectively measured proficiency.

The findings of this study provide several pedagogical implications for English language education in ESL-oriented contexts. First, language programs should move beyond an exclusive focus on linguistic competence and place greater emphasis on developing learners' confidence and willingness to communicate. Teachers should create supportive and low-anxiety environments where students can use English meaningfully without excessive concern about making mistakes. Second, because academic year was not significantly associated with readiness, English programs should integrate authentic communication opportunities across all years of study rather than assuming that readiness will develop naturally over time. Finally, since students with lower self-rated proficiency reported lower levels of readiness, additional support should be provided through scaffolded communicative activities, gradual exposure to real-world English use, and confidence-building strategies that encourage active participation in communication.

6. Conclusion

The findings suggest that English-major students generally perceived themselves as being ready for real-world English use. However, the study does not provide direct evidence regarding the effectiveness of the curriculum or instructional practices that may contribute to such perceptions. Within three dimensions, willingness to communicate emerged the most developed, suggesting that current pedagogical approach has included friendly environment where learners feel free to practice English. In addition, the difference across English proficiency levels highlights the need to pay attention to learners' linguistic capacity improvement because it is also one of main factor that affect their level of readiness. Strong positive correlations among communicative competence, willingness to communicate, and L2-self-confidence further confirm that improving learner readiness needs a combined activation of cognition, affection, and behavior.

The study has three limitations. First, most of the participants employed in this study are students in year 3 and year 4, while students in year 1 and year 2 remain fewer amounts. This unequal distribution might influence the ANOVA results. Second, despite the effort to recruit as many participants as possible in such context, the study was conducted in a single institution, which limits the generalizability of the findings. Therefore, while the findings provide valuable insights into learner readiness in a specific context, they might not fully capture the diversity of experiences across different institutions or learner groups. Third, the significant differences across proficiency levels should be interpreted with caution because the beginner and advanced groups consisted of only six and five participants, respectively, which may have reduced the stability of the statistical estimates.

Improving from current limitation, future research should enlarge the numbers of participants in different institutions to provide a more comprehensive understanding of learners' readiness toward English as second language using. In addition, a qualitative approach, such as interview or classroom observation, should be added to deeper exploration of learners' experiences, perceptions, and challenges in using English in real-world contexts. By combining a mixed-method approach and a larger research scope, future studies can develop holistic understanding and suggest more practical pedagogical implications for supporting learners in ESL-oriented context.

References

Bensalem, E., & Thompson, A. S. (2021). Multilingual effects on EFL learning: A comparison of foreign language anxiety and self-confidence experienced by bilingual and multilingual tertiary students. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2021.1943306

Chorrojprasert, L. (2020). Learner readiness – Why and how should they be ready? *LEARN Journal: Language Education and Acquisition Research Network, 13*(1), 268–274.

Clément, R. (1980). Ethnicity, contact and communicative competence in a second language. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 147–154). Pergamon Press. https://doi.org/10.1016/B978-0-08-024696-3.50027-2

Dang, T. H., Chau, T. H., & Tra, T. Q. (2020). A study on the difficulties in writing argumentative essays of English-majored sophomores at Tay Do University, Vietnam. *European Journal of English Language Teaching, 6*(1), 202–238. https://doi.org/10.46827/ejel.v6i1.3389

Delgado, M. R. (2021). How perception of control shapes decision making Kainan S Wang, Yun-Yen Yang 3 and Mauricio R Delgado 3. *Current Opinion in Behavioral Sciences*, *41*, 85-91.

Do, H. T. (2000). Foreign language education in Vietnam: The emergence of English and its impact on higher education. In J. Shaw, D. Lubeska, & M. Noullet (Eds.), *Language and development: Partnership and interaction: Proceedings of the Fourth International Conference on Language and Development* (pp. 29–42). Asian Institute of Technology.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course* (2nd ed.). Lawrence Erlbaum Associates.

Gilquin, G., & Granger, S. (2011). From EFL to ESL: Evidence from the International Corpus of Learner English. In J. Mukherjee & M. Hundt (Eds.), *Exploring second-language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 55–78). John Benjamins Publishing Company. https://doi.org/10.1075/scl.44.04gra

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.

Kachru, B. B. (1992). Models for non-native Englishes. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (pp. 48–74). University of Illinois Press.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a second language: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562.

McGroarty, M. (1984). Some meanings of communicative competence for second language students. *TESOL Quarterly, 18*(2), 257–272.

Nguyen, H. T. M., Nguyen, H. T., Nguyen, H. V., & Nguyen, T. T. T. (2018). Local challenges to global needs in English language education in Vietnam: The perspective of language policy and planning. In P. C. G. Lian, C. Chua, K. Taylor-Leech, & C. Williams (Eds.),

Un(intended) language planning in a globalising world: Multiple levels of players at work (pp. 214–233). De Gruyter Open. https://doi.org/10.1515/9783110518269-012

Prime Minister of Vietnam. (2025). *Decision No. 2371/QĐ-TTg dated October 27, 2025 on the approval of the project "Making English a second language in schools for the period 2025–2035, with a vision to 2045"*. https://vanban.chinhphu.vn/?pageid=27160&docid=215728

Savignon, S. J. (2018). Communicative competence. In J. I. Liontas, T. International Association, & M. DelliCarpini (Eds.), *The TESOL encyclopedia of English language teaching*. Wiley. https://doi.org/10.1002/9781118784235.eelt0047

Scarcella, R. C., Andersen, E. S., & Krashen, S. D. (1990). *Developing communicative competence in a second language*. Heinle & Heinle.

Schneider, E. W. (2003). The dynamics of New Englishes: From identity construction to dialect birth. *Language, 79*(2), 233–281.

Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.

Lightbown, P. M., & Spada, N. (1999). *How languages are learned* (2nd ed.). Oxford University Press.

Hoang, V. V. (2010). The current situation and issues of the teaching of English in Vietnam. *Ritsumeikan Studies in Language and Culture, 22*(1), 7–18. https://doi.org/10.34382/00002567

Tran, N. D., & Nguyen, T. L. (2023). Reading comprehension difficulties among English majored sophomores at a university in Vietnam. *International Journal of Scientific Research in Science and Technology, 10*(2), 890–905. https://doi.org/10.32628/IJSRST523102149

Tunçel, H. (2015). The relationship between self-confidence and learning Turkish as a foreign language. *Educational Research and Reviews, 10*(18), 2575–2589. https://doi.org/10.5897/ERR2015.2445

Vo, P. Q., Pham, T. M. N., & Ho, T. N. (2018). Challenges to speaking skills encountered by English-majored students: A story of one Vietnamese university in the Mekong Delta. *Can Tho University Journal of Science, 54*(5), 38–44. https://doi.org/10.22144/ctu.jen.2018.022

Bionote

Nguyen, Thi Thanh Loan is a full-time lecturer at the Faculty of Foreign Languages, HUFLIT University, Vietnam. She holds an MA in Education with a major in TESOL and has over 10 years of teaching experience. Her research interests include AI applications in language teaching and English as a Second Language (ESL). She has published in international conference proceedings on language research and teaching with AI. Her current work focuses on students' readiness for real-world English use in ESL-oriented contexts.

#144 ENGLISH PHARMACEUTICAL TERMINOLOGY AT LHU FACULTY OF PHARMACY: CURRENT USAGE, LEARNING DIFFICULTIES, AND PEDAGOGICAL IMPLICATIONS

Nguyen Huu Thinh[1], *Pham Huu Duc*[2]

Abstract

English pharmacology terms become increasingly significant as a means of preparing pharmacy students for both academic study and professional communication in modern health care environments. This research explores the way students at Lac Hong University (LHU) are exposed to and learn English pharmaceutical terms, their problems and difficulties and the methods of teaching which can help to overcome these difficulties. The research is based on a mixed-method approach with questionnaires being filled in by 80 students and semi-structured interviews being conducted with 10 participants (Babar, 2015). The quantitative analysis was based on such factors as exposure and usage of the language, knowledge level, difficulties in studying and difference between academic years, while qualitative analysis was applied to detect certain themes in the respondents' experiences. As for the survey findings, the exposure and usage level is moderate (M = 3.97), while knowledge and confidence are only average (M = 3.03). Learning difficulties are high (M = 4.25). In terms of exposure, the highest was found in the internships or professional observations (M = 4.55, SD = 0.50). On the other hand, from interviews, limited opportunities for practicing, continued dependence on translation, and more contextual teaching were highlighted. For academic years, there is no significant difference in competence in terms of exposure to pharmacy terminology, F (2, 77) = 0.903, p = .410, eta squared = .023. In addition, it is proposed that the teaching of pharmacy terminology should not remain at the level of theoretical exposure but proceed to the stage of practice in professional situations. Pedagogical implications include CLIL-based approach, case-based learning, pronunciation help, scaffolding of ESP module lessons, and appropriate assessments.

***Keywords*:** English pharmaceutical terminology, ESP, CLIL, pharmacy education, learning difficulties, mixed-methods research

1. Introduction

English is one of the principal languages used for science, medicine, and pharmacy purposes. Nowadays, pharmacy students find it necessary to study English textbooks and articles, to obtain drug information and guidelines in English (Aulton & Taylor, 2022; DiPiro et al., 2020; Murphy, 2022). Hence, it is not any more a choice of advantage to be proficient in the terminology of English pharmacy; it becomes essential to succeed at the academic level and prepare for work. Terminology plays an important role in such fields as reading, reporting, prescriptions, advising patients, and communicating with colleagues from other health care specialties. Mistakes in the use of technical terms may cause a lack of understanding, inaccurate communication, and even patient safety issues (Beardsley et al., 2020). The matter becomes more complicated for Vietnamese students in connection with the growing internationalization of health care education.

[1]*Nguyen Huu Thinh, Lac Hong University*
[2] *Pham Huu Duc, School of Languages, International University -- VNU HCMC*

In the case of Lac Hong University (LHU), pharmacy students come into contact with the English terminology via courses in the discipline, reference books, and internship assignments, facilitated by a context of program in which there is provision for professional training facilities (Faculty of Pharmacy, Lac Hong University, 2020). Nevertheless, being in contact with the terms does not mean that they will use them confidently and correctly. Quite a number of students have problems with comprehending, memorizing, pronouncing and using the terms properly.

1.1 Problem statement

While LHU pharmacists' students learn pharmaceutical English vocabulary in the classroom setting, they still have difficulty applying it in a consistent manner in various learning and practice contexts. The acquisition process seems to promote recognition of vocabulary items rather than their context-based use, particularly when the context involves dealing with patients and communicating with colleagues. The pedagogical challenge here is that students learn the vocabulary passively without building communicative skills required for using it in a professional setting (Beardsley et al., 2020; Rutter, 2021).

1.2 Research objectives

The research was carried out in order to investigate the use of pharmaceutical terminology in English by pharmacy students at LHU, to determine the problems associated with acquisition and use of pharmaceutical terminology by the students, to examine the students' perceptions of their acquisition and use of pharmaceutical terminology, and to make some pedagogical suggestions.

1.3 Research questions

The study addressed the following questions:

1) What difficulties do students encounter in learning and using pharmaceutical terminology in academic and practicum settings?

2) What are students' perceptions of learning and using English pharmaceutical terminology?

1.5 Delimitations

The research is limited to pharmacy students in the Faculty of Pharmacy at Lac Hong University. The research examined the use, comprehension, and difficulties in learning English pharmaceutical terminology. The results of the quantitative data collection were gathered from questionnaires while the results of the qualitative data collection were from interviews. Hence, the research is not meant to measure the level of competence in terminology or generalize across all pharmacy students in Vietnam.

2. Literature review

2.1 Pharmaceutical terminology

Pharmaceutical terminology means the specific linguistic framework employed in describing drugs, dosage forms, mode of action, methods of manufacture, pharmacokinetics, and applications of drugs in treatment (Aulton & Taylor, 2022; Arnold, 2020). Pharmaceutical terminology is normally characterized by precision, systematicity, and international use, but at the same time, it is challenging for students since it is characterized by complex vocabulary, which has its origins in the Latin and Greek languages (Chabner, 2023; Duy, 2025). Besides, most of the pharmaceutical terms are not only used as labels, but they also convey vital conceptual distinctions (Allen & McPherson, 2023; DiPiro et al., 2020).

Another problem is the complexity and accuracy of professional language. In pharmacy, even slight discrepancies in the use of terms can point to significant discrepancies in the route of administration, drug formulation, contraindications, or interactions (Preston, 2019; Aulton & Taylor, 2022). For international students who use English as their second language, the problem is intensified by the requirement to comprehend all these factors at once.

2.2 ESP and terminology learning

The English for Specific Purposes (ESP) focuses on the learning process in terms of its direct relation to the needs of learners both in academic and professional spheres. For pharmacy studies, ESP could assist in teaching technical language, not through mere memorizing wordlists, but through its application in real practices of reading, discussion, report writing, and other actions typical for pharmacists' profession (Babar, 2015; Beardsley et al., 2020). Also, an approach to learning through ESP pays much attention to the needs analysis, task authenticity, and communication purposes. For the case of pharmacy students, the target language domain would be specific for the discipline rather than everyday conversation (Babar, 2015). Thus, good ESP assistance would bridge the gap between lexical competence and actual language use through the practical application of the terminology.

2.3 CLIL and pharmacy education

CLIL is another effective approach that can be considered as it combines content and language learning. In pharmacy education, CLIL may comprise using English materials, bilingual teaching, or performing certain tasks which require processing and conveying of disciplinary content in English (Allen & McPherson, 2023; Mahato & Narang, 2018). It is quite appropriate for such situations when professional knowledge becomes available through international resources. CLIL is an effective pedagogical approach as it does not perceive language learning as a separate activity from content learning. On the contrary, it stimulates learners to acquire content and language at the same time. In case of pharmacy education, such an approach can help to promote more in-depth involvement in learning terminology as it will allow placing terms into the context of problem-solving and analysis (Allen & McPherson, 2023; Mahato & Narang, 2018).

2.4 Related studies

Research and professional literature have found that terms in the medical and pharmaceutical field tend to be challenging due to their length, morphology, pronunciation challenges, and contextual understanding (Chabner, 2023; Bonsell, 2024). In addition, they have pointed out that a theory-driven approach does not ensure that students will be adequately prepared for professional communication when it comes to terminology (Beardsley et al., 2020; Rutter, 2021). This means that the existing research highlights the common dichotomy between terminology recognition and use as an important challenge. Furthermore, it indicates that terminology acquisition is not only an issue of learning vocabulary but also a larger educational challenge that includes curriculum design, exposure, and practice opportunities (Babar, 2015; Beardsley et al., 2020). This approach is critical for the current study as it provides an understanding of students' challenges from different perspectives.

3. Methodology

3.1 Research design

This study used a mixed methods approach that involved the collection of both quantitative and qualitative data (Babar, 2015). The quantitative method was used for identifying trends in the usage of terminologies, their understanding, and their level of

difficulty, while the qualitative method was used to probe into the experiences of the learners and their explanations.

3.2 Participants

A total of 80 pharmacy students from the Faculty of Pharmacy, LHU were interviewed. Among them, 52 students were females and 28 were males. According to the academic year, 6 students belonged to Year 2, 53 students were from Year 3, and 21 students were from Year 4. Also, 10 interviewees were responsible for the interview data. Based on the interview file, one may assume that those people had some experience related to pharmacists and pharmacological language.

3.3 Research instruments

Two tools were employed in the research. A questionnaire consisting of 20 five-point Likert scale items investigated the frequency of exposure and use, level of understanding and confidence, learning difficulties, and preferred pedagogical assistance forms. Interviews in a semi-structured manner were conducted to explore the perceptions of participants about commonly used terms, contexts of their use, learning difficulties, opportunities to practice, teaching methods, and improving instructions. The questionnaire was divided into four major themes: frequency of exposure and use (Questions 1-5), self-perceived understanding and confidence (Questions 6-11), learning difficulties (Questions 12-17), and pedagogical preferences (Questions 18-20).

3.4 Reliability and validity considerations

Reliability and validity measures for the instrument need to be provided for the paper to be suitable for publication (Babar, 2015). Based on preliminary reliability estimates for the present dataset, the instrument demonstrates low values of internal consistency across several different groups of items. For instance, Cronbach's alpha is equal to 0.049 for all 20 items of the questionnaire, 0.109 for the five items that refer to exposure and use, -0.395 for the six items measuring understanding and confidence, 0.174 for the six items on the difficulty level, and -0.043 for the three items assessing pedagogical support. These numbers indicate that one should not treat the instrument as a reliable unified scale now. This matter should be addressed. If the paper is published in a journal in the future, one needs to demonstrate additional proof of content validity, piloting, and re-development of the items and the scales before making claims regarding the psychometric properties of the instrument.

3.5 Data analysis

Descriptive statistics such as mean, standard deviation, frequencies, and percentages were used for quantitative data analysis. An analysis was carried out comparing the students' year levels in the case of the self-reporting understanding/confidence composite variable. The one-way analysis of variance technique was used to determine whether there is a significant difference in self-reporting competencies in terms of academic years. Thematic analysis was used for qualitative data. As opposed to coding all utterances in the qualitative interview process, the analysis focused on recurring themes concerning memorizing, pronunciation, translation, practice, reading, speaking, internship, and suggestions for improvement.

4. Findings

4.1 Difficulties students encounter in learning and using pharmaceutical terminology

The survey reveals four key trends. First, the participants claimed a fairly high level of exposure and usage of English pharmacological terms (M = 3.97). Second, self-assessed levels

of comprehension and mastery of the knowledge were quite average (M = 3.03). Third, the difficulties of acquiring knowledge were very high (M = 4.25). Fourth, there was a high demand for pedagogical assistance (M = 4.21). It can be concluded that frequent encounters with terminology do not guarantee comprehension and proficiency in it.

4.2 Frequency and contexts of terminology use

Students reported different levels of exposure in different contexts, as can be seen in Table 1.

Table 1

Contexts of terminology exposure and use (N = 80)

Context	M	SD
Specific subjects	3.61	0.49
Reading textbooks or study materials	3.54	1.01
Practical courses or laboratories	3.56	0.87
Internship or professional observation	4.55	0.50

The highest mean score was observed in internship or professional practice (M = 4.55, SD = 0.50). In this respect, it can be concluded that the exposure to English drug terminology is most intense in a situation where students are getting nearer to professional practices. At the same time, the mean scores for other academic disciplines, books, and practical disciplines were less intense. Therefore, one can assume that terminology becomes particularly relevant when students are dealing with professional activities.

4.3 Self-reported understanding and confidence

The students' level of understanding and confidence lagged behind the amount of exposure that they had received. The composite of the six items on understanding and confidence was classified into broad bands of performance as shown in Table 2.

Table 2

Distribution of self-reported understanding and confidence

Level	Frequency	Percentage
Very poor	0	0.0
Poor	4	5.0
Average	72	90.0
Good	4	5.0
Very good	0	0.0

The highest percentage of students belonged to the average category (90%), and only 5% belonged to the 'good' category and 5% belonged to the 'poor' category. None of the respondents belong to the category of 'very good' and 'very poor'. It is evident that there is a concentration towards the middle category: students have knowledge about the terms used in pharmaceuticals, but very few of them think they are proficient or not proficient.

4.4 Learning difficulties

The survey identified several major difficulties, as this can be seen in Table 3.

Table 3

Main learning difficulties

Type of difficulty	M	SD	Rank
Pronunciation	4.51	0.50	1
Using terms in practical contexts	4.42	0.50	2
Memorizing meaning	4.36	0.66	3
English background affects learning	4.24	0.68	4
Distinguishing similar terms	4.21	0.76	5
Limited access to specialized literature	3.76	0.92	6

Pronunciation was identified as the top rated difficulty (M= 4.51, SD = 0.50) followed closely using the terminology in practical situations (M= 4.42, SD = 0.50) and memorization of meaning (M= 4.36, SD = 0.66). This indicates that the difficulty does not lie just in the identification of the technical terminologies when presented in written form.

4.5 Comparison across academic years

A comparison of the understanding/confidence composite in different academic years shows only small differences, as can be seen in Table 4.

Table 4

Self-reported understanding and confidence by academic year

Academic year	N	M	SD
Year 2	6	3.11	0.36
Year 3	53	3.00	0.30
Year 4	21	3.08	0.24

This descriptive pattern reveals very little difference in values from year to year, where Year 2 students have a mean value of 3.11, Year 3 students have 3.00, and Year 4 students have 3.08. To investigate if these differences were significant, a one-way ANOVA test was performed, which is shown in Table 5.

Table 5

ANOVA for self-reported understanding and confidence by academic year

Source	SS	df	MS	F	p	eta squared
Between groups	0.148	2	0.074	0.903	.410	0.023
Within groups	6.321	77	0.082	—	—	—
Total	6.469	79	—	—	—	—

ANOVA was not found to be statistically significant with F (2, 77) = 0.903, p = .410. The effect size was small (eta squared = 0.023). Consequently, the working hypothesis

predicting an association between the students' academic year and their perceived terminological competence could not be supported by the data set used.

4.6 Qualitative themes from interviews

The interview data corroborated and expanded upon the survey data. A number of recurring themes emerged from the interview responses. The first theme was the importance of practice. Practice-related terminology was very frequent in the interview corpus, which indicates that the lack of practice opportunity is an important issue. The second theme was memorization and remembering, which is consistent with the survey data that indicated remembering meaning to be an important difficulty. The third recurring difficulty from the interviews was pronunciation, which corroborates the quantitative result that pronunciation is the greatest difficulty among all difficulties. The fourth issue that emerged from the interviews was the prevalence of references to reading compared to the rare mentions of speaking, which indicates that vocabulary is usually learned receptively rather than interactively. Finally, translation and bilingual terminology were mentioned as an issue multiple times, corroborating the survey results that indicate that the use of English - Vietnamese terminology helps both students and teachers learn new words. Overall, these themes prove that the terminology learning difficulties are not merely lexical in nature, but also pedagogical and situational

5. Discussion

For addressing the first research question, the results point to rather high exposure to the English pharmacy vocabulary used in practice by the participants of the study. At the same time, this frequency is not accompanied by the same level of knowledge and competence. This disparity between the relatively high exposure and use mean (3.97) and considerably lower understanding and confidence mean (3.03) implies that the students perceive vocabulary more easily than they understand, utilize and work with it.

It is significant since pharmacy requires not only awareness of the written terminology but also its correct pronunciation, distinguishing of the related concepts, and proper usage in communication. According to the results of this study, pronunciation, applicability, and memorizing are the top three difficult areas in learning pharmacy terminology. It means that terminology acquisition in this field includes linguistic, conceptual, and communicative aspects (Chabner, 2023; Beardsley et al., 2020). Year-level data was equally illuminating. Early versions of the manuscript assumed that students at higher levels of academic year would perceive themselves as more competent owing to accumulated experience. The present data set, on the other hand, indicates small and non-statistically significant differences between different years. It should be noted that this does not imply that experience is unimportant. What it implies is that the year level itself cannot necessarily lead to greater perceptions of competence.

Regarding the second research question, one can understand how the interview responses relate to this finding. Students often mentioned lack of practice, dependence on translations, and preference of reading over speaking. Such comments suggest that, despite increased exposure to terms, the process of translating their passive knowledge into active professional usage is not adequately facilitated. It could be used to explain why self-assessment of proficiency was found to be centered around the mean value. The other significant implication relates to the method used for assessment. The poor reliability estimates indicate that questionnaire measures several different constructs, not a stable one. Thus, the results cannot be generalized; however, the instrument can still be used as a descriptive measure.

6. Pedagogical implications

These results have some practical implications on how to teach and design curricula in connection with Targeting pronunciation in a systematic way, increasing realistic practice opportunities, moving away from rote learning techniques, creating scaffolding of ESP, enhancing CLIL-based integration, and assessment aligned with communicative usage.

As far as targeting pronunciation goes, pronunciation is the most highly rated difficulty, so term teaching has to include pronouncing terms systematically and not as an addition. It can involve marking syllables, practicing stress, repeated pronunciation, teacher modeling, sound support, and quick oral check-ups. In the field of pharmacy, the ability to pronounce things correctly is not only language skills but also professionalism (Beardsley et al., 2020).

Regarding using terms in realistic contexts, the application of terms in practical situations turned out to be the second greatest challenge, thus instructors should create assignments where learners have to use the terms in real problem-solving. Discussions of cases, interpreting prescriptions, small role-plays, counselling activities and the interpretation of interactions between medications may be used to bridge the gap between memorization and applying terminology (DiPiro et al., 2020; Preston, 2019; Rutter, 2021).

To move away from memorization approaches only, the results of interviews and surveys show that memorization-based learning processes still dominate in terminology learning. Even though memorization techniques can be helpful when learning technical vocabulary, it should be combined with other skills such as morphological analysis, contextual usage, practicing retrieving and reusing terminology (Chabner, 2023; Duy, 2025).

To provide scaffolding of ESP, the development of the English for Pharmacy module or ESP strand can bring together repeating terms of pharmacology, pharmaceutics, clinical pharmacy, and professional communication in internships. The development of ESP should involve recognition, controlled practice, and performance of simulations in professional practice (Babar, 2015; Beardsley et al., 2020).

For promoting CLIL-based integration, the high frequency of terminology in internships and observations demonstrates that terminology gets meaning only in contextually enriched environment. In this regard, CLIL-based teaching might be helpful in working with pharmaceutical notions and language at once. Subject lecturers might incorporate bilingual glossaries, terminological case tasks, and content-based tasks in English into specialized subjects (Allen & McPherson, 2023; Mahato & Narang, 2018).

In terms of assessment alignment with communicative practice, the assessment is focused mainly on the recall of terms, and students will continue treating terminology learning as the memorization process. At the same time, if the assessment will include explanation, oral reporting, case interpretation, and reformulation of information in patient-friendly language, students might start using terminology productively (Babar, 2015; Beardsley et al., 2020).

7. Conclusion

From this research, it can be seen that the knowledge of the English terminology related to pharmaceuticals is an important component of the pharmacy curriculum at LHU; however, the problem of difficulties in comprehending, remembering, pronouncing, and using specialized terminology remains serious.

7.1 Summary of the main findings

Moreover, it is evident from the findings that the pronunciation, application, and memorization of specialized terms are the biggest obstacles. It must also be noted that the

expected year-level effect was not revealed statistically in the current dataset. Consequently, it can be argued that the proposed paradigm shift towards more practical instruction of terminology is justified. In particular, the following methods seem especially relevant: pronunciation assistance, case approach, scaffolding of ESP, CLIL, assessment.

7.2 Limitations

The following are some of the limitations that need to be taken into consideration. The first is that the study was undertaken within one organization only, hence limiting its generalizability. The second is that the study results are based on self-report data rather than performance measurement. The third is that the sample used for interviews was quite small and should not be considered as representative but as an illustration. The fourth is the lack of reliability of the questionnaire. In other words, the reliability of the questionnaire was quite low. Therefore, the survey can only be viewed as a descriptive instrument rather than as a scale.

7.3 Future research

The studies should consider expanding the investigation to multiple settings and comparing programs with various degrees of English integration. The studies should also incorporate quantitative data on terminological competence, which may take the form of vocabulary tests, pronunciation tests, role-playing exercises, or other performance measures. An experimental study or a quasi-experimental study will be particularly relevant when testing the impact of ESP modules, CLIL-inspired integration, case-based learning, and simulation training on the usage of terminological units (Babar, 2015; Bonsell, 2024). The subsequent studies will need to reconsider and validate the questionnaire used. Through item revision, piloting, and factor analysis, the researchers will be able to ascertain whether terminology exposure, terminology understanding, difficulty of terminological units, and support needs should be considered separate dimensions.

References

Allen, L. V., & McPherson, T. B. (2023). *Ansel's pharmaceutical dosage forms and drug delivery systems* (10th ed.). Wolters Kluwer.

Arnold, R. J. G. (2020). *Pharmacoeconomics: From theory to practice* (2nd ed.). CRC Press.

Aulton, M. E., & Taylor, K. M. G. (2022). *Aulton's pharmaceutics: The design and manufacture of medicines* (9th ed.). Elsevier.

Babar, Z. U. (2015). *Pharmacy practice research methods*. Springer.

Beardsley, R. S., Skelton, M. Z., & Kimberlin, C. L. (2020). *Communication skills in pharmacy practice: A practical guide for students and practitioners* (7th ed.). Wolters Kluwer.

Bonsell, R. (2024). *Pharmacy technician certification exam: Practice question workbook*. Coventry House Publishing.

Chabner, D.-E. (2023). *Medical terminology: A short course* (9th ed.). Elsevier.

DiPiro, J. T., Yee, G. C., Posey, L. M., et al. (2020). *Pharmacotherapy: A pathophysiologic approach* (11th ed.). McGraw-Hill.

Duy, N. T. (2025). *Medical terminology* [in Vietnamese]. Women's Publishing House.

Faculty of Pharmacy, Lac Hong University. (2020). *Faculty profile*.

Mahato, R. I., & Narang, A. S. (2018). *Pharmaceutical dosage forms and drug delivery* (3rd ed.). CRC Press.

Murphy, J. E. (2022). *Clinical pharmacokinetics* (7th ed.). American Society of Health-System Pharmacists.

Preston, C. L. (Ed.). (2019). *Stockley's drug interactions* (12th ed.). Pharmaceutical Press.

Rutter, P. (2021). *Community pharmacy: Symptoms, diagnosis and treatment* (5th ed.). Elsevier.

#146 EXPLORING THE EFFECTS OF SPEAKEASY AI APPLICATION ON SPEAKING FLUENCY AMONG FIRST-YEAR ENGLISH-MAJOR STUDENTS AT HUFLIT

Nguyen Thi Huyen Tran[1]

Abstract

AI-powered speaking applications have been gaining more traction in recent years. They are more engaging than traditional ones because they usually come with interactive elements designed to grab learners' attention and motivate them to practice speaking consistently. As such, many teachers have started using tech apps as teaching aids. This research aims to assess the effectiveness of SpeakEasy AI on students' speaking development. SpeakEasy AI is an AI-powered web-based application that allows learners of English to practice speaking through various modes – chatting in real-time with the bot, taking part in speaking simulation tests, and uploading speech audio recordings to get speaking scores. A mixed-method experimental approach was adopted, involving 80 first-year English-major students from HUFLIT who were majoring in English. The experimental group was asked to use SpeakEasy for 12 weeks, while the control group was taught with conventional teaching. Both groups took pre-tests and post-tests and answered questionnaires. The application's effects were measured via statistical analysis and compared with those of traditional teaching methods. Several students were also interviewed to obtain qualitative data. The results showed that students who used SpeakEasy significantly outperformed those in the control group in terms of fluency and overall speaking proficiency. The students who used the app also reported positive attitudes towards the application, especially speaking anxiety reduction and increased confidence. Data from the interview suggested other benefits in addition to improved speaking performance and provided some insights into improving the application. More studies are needed to determine the long-term effect of the application on different populations of EFL learners.

Keywords: SpeakEasy AI app, AI-powered chatbot, speaking fluency, speaking competency, EFL learners.

[1]*Nguyen Thi Huyen Tran - Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT)*
Email: trannth1@huflit.edu.vn

1. Introduction

1.1 Background and Statement of the Problem

Speaking ability is arguably one of the most crucial components in mastering English as a second language. Not only does speaking underlie most language skills and enable learners to communicate, but it also builds confidence and social skills in students that allow them to thrive in virtually every situation (Baker & MacIntyre, 2000). However, many learners feel uncomfortable and experience anxiety when speaking English. According to Amoah and Yeboah (2021), students hesitate to speak in class because they are afraid of making mistakes that may make them look bad in front of others. They are also worried that their classmates or teachers might give negative feedback. Engeström (2015) suggested that a person's emotional state can influence their ability to express ideas clearly. Feelings such as insecurity, shyness, nervousness, and worry often appear and make students interrupted utterances. He also revealed that another factor contributing to speaking challenges is learners' limited exposure to English-speaking environments as they have few opportunities to practice the language in authentic communicative contexts.

Teachers therefore play an important role in creating a supportive language environment where students feel secure to practice speaking. They can encourage students to speak both inside and outside the classroom by providing different language situations and social contexts. In addition, AI-powered applications can assist students by giving them more opportunities to practice communication. Artificial Intelligence (AI) technologies have been implemented into educational classrooms and have become useful tools to improve speaking proficiency within second language acquisition, such as English as a Foreign Language (EFL). The implementation of AI into EFL can be seen through tutoring systems, speech recognition applications, and virtual conversation partners. By allowing immediate feedback, realistic conversations, and personalized learning, these tools can help overcome existing issues as well as help cultivate oral skills.

In recent years, several applications for practicing English speaking skills have entered the market. The majority of them are based on Automatic Speech Recognition technology (ASR). ASR can provide instant feedback to learners regarding pronunciation and speaking performance. This technology became one of the most frequently used approaches in CALL due to its instant feedback system (Liakin, Cardoso, & Liakina, 2017; McCrocklin, 2016). Applications like ELSA Speak or Duolingo are some of the most popular applications for language learners. However, each application focuses on a particular area of language learning. For example, ELSA Speak has been widely used as an AI-based tool designed primarily to improve learners' pronunciation accuracy through automatic feedback (Karim et al., 2023; Kumalasari & Wulandari, 2025). By contrast, Duolingo focuses more on vocabulary development and basic grammar practice through short, gamified learning activities that encourage repeated practice and learner engagement (Loewen et al., 2019). Although ASR technology allows students to detect and correct pronunciation mistakes, meaningful conversational practice with AI assistants is often only available in premium paid versions. This creates a practical limitation for teachers who want to apply these tools in large classes with diverse groups of students, as not all learners can afford paid subscriptions.

In addition, although speaking chatbots have given way to promoting speaking practice among students because they can create a personalized learning environment (Jeon, 2024), they struggle to simultaneously function as both interactive conversational agents and reliable assessment tools where students can experience standardized test environment and receive comprehensive evaluation. Specifically, they do not offer students a centralized environment

in which to complete IELTS-related speaking activities, take the IELTS speaking test, and receive scores based on defined assessment criteria. Students, therefore, have to utilize different applications for IELTS speaking practice and assessment.

SpeakEasy AI was created to overcome these drawbacks. The application aims to simulate, support practice, and evaluate English speaking skills. SpeakEasy AI accomplishes this goal by delivering an AI-driven web-based platform that offers real-time speaking practice and feedback. This allows users to have a genuine speaking experience. SpeakEasy AI can also simulate speaking test environments to help learners prepare for speaking exams and assess them using standard rubrics such as IELTS, VSTEP or custom rubrics.

1.2 Research Objectives and Questions

The purpose of this study is to gain a better understanding of the effectiveness and user experience of the SpeakEasy AI application as a speaking instruction tool through the lens of students' fluency ability in speaking performance and their perspectives.

To achieve the purpose of the study, the following research questions were formulated:

1. To what extent does the SpeakEasy application improve students' speaking fluency?

2. What are the students' perceptions of using the SpeakEasy application in learning speaking skills?

1.3 Significance of the study

This study offers an insight into the effects of integrating an AI-powered application on enhancing the speaking fluency of English-major students in the Vietnamese university context. The study may also provide viable suggestions for educators on how to inccorporate an AI application as a supplementary tool for English speaking practice and test preparation. In addition, the findings of this study will contribute to the growing body of research supporting the integration of AI-powered applications into EFL speaking instruction.

1.4 Scope of the Study

This tudy was limited to examining the effectiveness of SpeakEasy AI application in improving the speaking fluency of first-year English-major students at HUFLIT. It focused on one AI-powered speaking application, one speaking criterion (Fluency and Coherence), one institutional context and over a 12-week intervention. Therefore, the results may not be applicable to students from other academic majors and educational institutions.

2. Literature Review

2.1 Speaking Fluency

Segalowitz (2010) defines speaking fluency as a three-component model consisting of cognitive fluency, utterance fluency, and perceived fluency. The first aspect, cognitive fluency, refers to how efficient speakers can access and process language, involving accessing vocabulary, encoding grammar, and producing sounds for speech. Compared to native speakers, second language (L2) learners' cognitive fluency is less developed because this process is controlled instead of being automatic. The increased cognitive effort leads to greater self-monitoring, slower speech and more frequent pauses. The second aspect, utterance fluency, is a feature that is observable and measurable. It encompasses the number of words or syllables articulated per minute, the quantity of syllables produced between pauses, and how often the speakers pause, hesitate or correct themselves. Cognitive fluency and Utterance fluency can enhance if language proficiency develops. In other words, if speakers have a large

pool of vocabulary, more solid grammar, and have a frequent use of them, they can retrieve the vocabulary and grammar faster and thus limit the frequency and the length of pauses (Tavakoli et al., 2020). Perceived fluency is related to the subjective assessment of listeners about a speaker's fluency, basing on the smooth, natural, and effortless level of the speech.

Some factors which adversely affect speaking fluency include limited linguistic ability such as insufficient vocabulary and grammar, infrequent listening and speaking practice, and a tendency to think in the native language while speaking (Alaraj, 2017). In addition, other affective factors such as the level of anxiety, motivation and learners' attitudes influence how effectively learners express themselves in real conversations. In other words, students' fixed mindset about their language capability and previous negative experiences may discourage them from actively participating in speaking activities (Bárkányi, 2021). In addition, the anxiety in speaking situations can interfere with the ability of learners to retrieve vocabulary automatically and may lead to excessive self-monitoring, thereby impeding cognitive fluency. Pabro-Maquidato (2021) also indicates that heightened anxiety is associated with increased pauses, hesitation, and errors, leading to reduced utterance fluency. Therefore, encouraging a growth mindset and establishing a supportive classroom atmosphere can help alleviate this fear. Brown (2017) highlights that improving fluency requires a supportive and low-pressure environment where learners are exposed to familiar topics, focus on meaning rather than form, and are encouraged to speak at a faster pace.

2.2 The application of AI in the field of EFL speaking instruction

A number of studies have been conducted to investigate the potentials and impacts of AI-powered tools on improving speaking skill. Qiao and Zhao (2023) examined the effects of AI-based instruction integrated with Duolingo application on four main components of speaking skills using IELTS speaking band descriptors: fluency, vocabulary, accuracy, and pronunciation and also the impact on self-regulation. The study was implemented with 93 Chinese EFL students of intermediate level during 13-week course using experimental design and pre- and post-test using standardized IELTS test. The experimental group received AI-based instruction with Duolingo application while the control group attended a traditional speaking course. Results showed that the experimental group performed better than the control group on four components of L2 speaking skills, along with self-regulation. As students practiced speaking repeatedly and reviewed vocabulary, they were able to remember words faster and construct grammatical forms more efficiently.

Celik et al. (2025) examined a very similar premise as they sought to understand if ChatGPT could be used as a virtual speaking tutor for EFL learners to improve speaking performance and self-efficacy. Forty-four sophomore university students enrolled in an advanced English course participated in the study and were randomly assigned to an experimental and control group. Whereas students in the control group received traditional classroom instruction, students in the experimental group practiced speaking with ChatGPT's voice-based conversation feature inside and outside of class. Results indicated that the experimental group significantly outperformed the control group on both speaking performance and confidence in speaking abilities. Additionally, learners shared positive perceptions about the experience, specifically noting flexibility, instant feedback, and repeated practice as benefits of using ChatGPT as a speaking tutor. However, some learners noted that ChatGPT did not offer the emotional support many human teachers provide.

Kang (2022) conducted a study of the impacts of Replika - an AI software offering oral-visual interaction with a 3D avatar on the speaking ability and affective characteristics of 61 ESL learners in an American college. The study employed a mixed-methods approach, in

which quantitative data were collected from the speaking tests and questionnaires, and qualitative data were obtained from the in-depth interviews. The findings revealed that both the experimental group using a paid version of the app and the control group interacting with native speakers achieved speaking enhancement, but the former showed more improvement on accuracy, fluency and coherence. Furthermore, learning via the app was found to be more beneficial for low-level students while practicing with native speakers was more effective for the high-level ones. In terms of affective aspects, the low-level students in the EG reported reduced stress and anxiety, as well as increased motivation, all of which were not observed by those in the CG.

Another study by Azizimajd (2023) also examined the effectiveness of Replika app into speaking activities on speaking fluency and Foreign Language Speaking Anxiety of 60 Iranian students of upper-intermediate level in 6 weeks. Experimental design using one-way and two-way ANCOVA was conducted, showing that students practicing with Replika surpassed the conventional class in terms of speaking fluency. In particular, they could produce longer utterances, reduce hesitation and pauses thanks to the tool's interactive and repetitive design. Also, they became less anxious in English speaking.

2.3 Research Gap

Taken together, the existing literature on the integration of AI-powered applications in speaking instruction provides valuable insights into their roles in enhancing many aspects of speaking ablility. Common features such as immediate and customized feedback, interactive learning experiences, real-time interaction and continuous practice have made significant contributions to improving spoken communication skill. In particular, AI-driven applications have been found to develop cognitive fluency (Qiao and Zhao, 2023) and utterance fluency (Azizimajd, 2023). Besides, these tools have been reported to effectively reduce stress and anxiety while boosting learners' motivation and confidence, thereby enabling them to communicate with ease and express ideas comfortably (Kang, 2022; Azizimajd, 2023; Celik et al., 2025). However, learning speaking through AI-based applications proved to to be more effective for low-level students than for the high-level ones (Kang, 2022). It was also observed that AI-driven tools does not provide the same level of emotional support as human teachers and therefore cannot substitute teachers in the language learning process (Celik et al., 2025).

Taking the aformentioned findings into consideration, the existing research on AI-based applications has shown some gaps that need to be tackled. First, while their findings have mainly focused on the overall effectiveness of AI applications in improving speaking ability, little attention has been paid to their impacts on specific sub-component such as fluency. Secondly, the use of interactive conversational feature in many platforms is often restricted to paid versions, making the implementation in large-scale classroom impractical. Another limitation is the lack of IELTS standardized testing environments where learners take IELTS simulated speaking test and receive criteria-based feedback. As a result, learners have to rely on separate platforms for speaking practice and assessment, highlighting the need for the provision of all-in-one solutions. Furthermore, studies investigating the effectiveness of AI-powered speaking applications on speaking fluency among English-major Vietnamese students in university context are still limited, especially those employing a mixed-method approach to explore students' perceptions through both quantitative and qualitative methods.

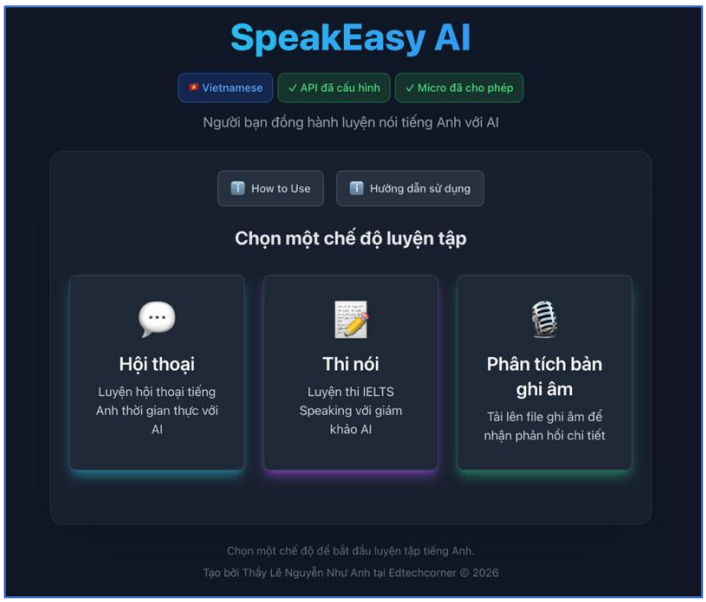
In response to the identified gaps and given the challenges that the researcher encounters in teaching speaking skill in the current context, this study aims to examine the effectiveness of SpeakEasy AI in enhancing speaking fluency and to capture students' subjective perspectives towards using this technological tool in their English learning.

2.4 SpeakEasy AI application

SpeakEasy AI is a web-based application developed as a part of Edtech Corner, an initiative that aims to bring unique technological educational tools to aid in English Language Teaching and Learning. It utilizes Google GenAI and Gemini 2.5 Flash Model and incorporates native-audio-preview using the live API, which allows audio to be sent and received natively without requiring it to go through text first. SpeakEasy AI aims to provide learners with AI-powered practice, testing and evaluation all in one place. To keep things accessible and affordable to learners, SpeakEasy AI was designed to be free and serverless. All a user needs to do is input their API key (ex. Gemini API).

Figure 1

Screenshot of SpeakEasy AI application



Real-time communication is one of the fundamentals behind SpeakEasy AI. Students can converse with AI-driven agents who provide immediate response, then they receive instant scores and feedback. Evaluation metrics can be preselected based on industry standard frameworks like IELTS or VSTEP or even customized rubrics. Another way to practice with SpeakEasy is through the “Simulated Test” mode, where learners can practice under test-like conditions that mimic conditions for IELTS, VSTEP or customized classroom assessments. This helps students get comfortable with the various test formats and helps them learn how to pace themselves. Learners can also control the difficulty of conversation topics and choose from different AI voices

3.Research Methodology

3.1 Pedagogic settings, Participants and Instrumentation

The research was conducted at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). The questionnaire survey involved 80 first-year English-major students enrolled in English Listening and Speaking courses. These classes were assigned by the Department of Foreign Languages during Semester 1 of the 2025–2026

academic year. One class served as the experimental group and the other as the control group, with 40 students in each. Both groups followed the same course structure, attending weekly sessions lasting three hours over a 12-week period.

In class, students received topic-based vocabulary words, focused grammar structures, and speaking instruction. Students were then given homework to practice speaking questions taught in class as well as a model IELTS speaking question on that topic. Students in the experimental group practiced speaking outside of class using the SpeakEasy AI app. They completed speaking activities on the app that were automatically scored and given feedback based on IELTS speaking criteria. Students submitted notes of their conversation with the chatbot as well as generated feedback from the app as a Word doc through Moodle for the instructor to track their progress outside of class. Students in the control group took a traditional approach to practicing and assessing speaking. Students recorded themselves speaking and submitted their spoken responses as audio files on Moodle for the instructor to assess using IELTS-based marking criteria..

3.2 Research design

The study employed a quasi-experimental mixed-methods design, combining quantitative evidence from pre- and post-tests with qualitative data from questionnaires and interviews to examine both the impacts of SpeakEasy AI on students' speaking fluency and their perceptions of the tool.

3.3 Data collection

3.3.1 Pretest and Posttest

The study employed a pre-test and post-test to measure the effectiveness of the intervention, using IELTS speaking assessment. The speaking questions in the pre-test and post-test were selected from the Cambridge Practice Test. Students' performances were evaluated based on the official IELTS speaking band descriptor (see Appendix 1) across four criteria: fluency and coherence, lexical resources, grammatical range and accuracy and pronunciation. For the purpose of statistical analysis, the scores related to "fluency and coherence" criterion were extracted.

The pretest was implemented after students had completed their first speaking lesson in order to minimize potential bias caused by students' unfamiliarity with test format. The posttest was conducted on the last day of the course. To enhance scoring reliability, all responses were recorded and subsequently assessed by a second rater - another lecturer who is also an experienced IELTS instructor. The Correlation between the scores given by the two raters are shown in the following tables.

Table 1

Correlation between pre-test scores given by two raters (n = 80)

	Rater 1	**Rater 2**	**Sig. (2-tailed)**
Rater 1	1	0.907	< 0.001
Rater 2	0.907	1	

Table 2

Correlation between post-test scores given by two raters (n = 80)

	Rater 1	Rater 2	Sig. (2-tailed)
Rater 1	1	0.927	< 0.001
Rater 2	0.927	1	

The Pearson correlation analyses showed that the correlation coefficients for both the pre-test and post-test scores assigned by the two raters exceeded 0.9 ($p < .001$). These findings indicate a strong positive relationship, reflecting a high degree of consistency in the raters' scoring. Therefore, the averaged scores from the two raters were considered reliable and objective for further statistical analyses, minimizing the potential influence of differences in rater scores.

3.3.2 Questionnaires

To explore students' perspectives on SpeakEasy AI, a questionnaire was developed using Google Form. It was structured into three section: Section 1 - Demographic information, Section 2 -Students' Perspectives on the effectiveness of SpeakEasy AI, Section 3 - Students' motivation when using SpeakEasy AI. The survey was implemented online, and its internal consistency was evaluated using Cronbach's alpha, as shown in the following table.

Table 3

Reliability of the Questionnaire (n = 40)

Part	Number of items	Cronbach's Alpha
Part II: Effectiveness of SpeakEasy AI	12	0.89
Part III: Motivation	6	0.86

The results of the reliability analysis indicates that all Cronbach's alpha coefficients exceeded 0.70. This demonstrated satisfactory reliability and strong internal consistency of the questionnaire, ensuring the suitability of the instrument used in the study.

3.3.3 Interview

The aim of the interviews was to examine learners' attitudes toward using SpeakEasy AI application for speaking practice, especially its effects on fluency. A structured format was used, in which all participants responded to the same set of questions in a consistent order. Five students were randomly chosen and labeled S1 to S5, and individual online interviews were conducted via Microsoft Teams to ensure privacy and encourage open sharing.

4. Findings

4.1 Fluency and Coherence Pretest and Posttest

4.1.1 Independent sample t-test of pretest results.

A descriptive statistics and an Independent samples t-test were employed to analyze the pre-test scores of the criterion "fluency and coherence" between the the control group (CG) and experimental group (EG) as presented in the following table. Before conducting the analysis, the assumption of normality was checked through Normal Q-Q plots. The plots indicated that the data from both groups were approximately normally distributed, suggesting that the requirement for parametric testing was satisfied.

Table 4

Descriptive Statistics of the Groups (n = 80)

Group	Sample size	Mean ± Standard Deviation
CG	40	5.03 ± 0.48
EG	40	4.89 ± 0.4

Table 5

Independent t-test results for Pre-test scores between the two groups (n = 80)

	Mean Difference	t	df	p (2-tailed) (sig.	95% CI
Equal variances assumed	0.13	1.37	78	0.17	(-0.06; 0.33)

** Levene's test for equality of variances: F = 0.42, p (sig.) = 0.51 (> 0.05), variances are assumed equal).*

The Levene's test produced a p-value of 0.51, which was higher than 0.05, indicating that the score variances of the two groups were similar. Therefore, the assumption of equal variances for the independent samples t-test was satisfied.

Independent samples t-test resulted in a t-score of 1.37 with 78 degrees of freedom and a p-value of .17. Because the p-value was greater than .05, there was no statistically significant difference between the experimental group's pre-test scores (M = 4.89, SD = 0.40) and the control group's pre-test scores (M = 5.03, SD = 0.48). Therefore, students in each group were similarly fluent and coherent prior to the intervention.

4.1.2 Independent sample t-test of posttest results.

An independent samples t-test was conducted for comparing EG and CG based on their post-test scores in the "fluency and coherence" criterion to check the efficacy of the proposed method. Normal Q-Q plots were employed to verify the normality assumption prior to running the statistical analysis. The results indicated that the score distributions were approximately normal in both groups which allowed parametric tests to be conducted.

Table 6

Descriptive Statistics of the Groups (n = 80)

Group	Sample size	Mean ± Standard Deviation
CG	40	5.48 ± 0.46
EG	40	5.86 ± 0.49

Table 7

Independent t-test results for Post-test Scores between the two groups (n = 80)

	Mean Difference	t	df	p (sig. (2-tailed)	95% CI
Equal variances not assumed	-0.38	-3.5	78	<0.001	(-0.59, -0.16)

** Levene's test for equality of variances: F = 1.46, p (sig.) =0.23>0.05, variances are assumed equal).*

Levene's test for equality of variances resulted in an F value of 1.46 and a p-value greater than 0.05. This result indicates that the variances of scores between both groups were equal. Thus, the equal-variance assumption was met.

Results from independent samples t-test revealed that there was a significant difference between post-test scores of students in the experimental group compared to students in the control group ($t = -3.5$, $df = 78$, $p < .001$). Specifically, students who learned with the SpeakEasy AI application had significantly higher scores compared to students who learned with the traditional teaching method (mean difference = -0.38, 95% CI: -0.59 to -0.16). The results imply that students from the experimental group significantly improved their learning performance. In conclusion, SpeakEasy AI application had a positive effect on students' fluency and coherence and outperformed the traditional teaching method.

4.2 Questionnaires

Forty students in the experimental condition participated in an online questionnaire survey about their experience of using SpeakEasy AI application. 47.5% were female participants and 52.5% were males. Table 8 shows the mean scores and standard deviation (SD) for each item in the questionnaire.

Table 8

Students' perceptions of using SpeakEasy AI (n = 40)

EFFECTIVENESS OF SPEAKEASY AI APPLICATION	MEAN	SD
After using the SpeakEasy AI app, I can speak more fluently.	4.08	0.35
I can retrieve vocabulary and grammar more quickly when speaking	3.9	0.44
The frequency of my pauses and hesitations (e.g., "uhm", "ah") has decreased.	3.93	0.41
The duration of my pauses (if any) has decreased.	4.00	0.39
The frequency of my self-corrections has decreased.	4.03	0.48
I can express my ideas more clearly without stopping in the middle.	4.00	0.39
I can use linking words and discourse markers appropriately in my speech.	4.13	0.46
The SpeakEasy AI app helps improve my score in the "Lexical Resource" criterion.	4.03	0.53
The SpeakEasy AI app helps improve my score in the "Grammatical Range and Accuracy" criterion.	3.83	0.44
The SpeakEasy AI app helps improve my score in the "Pronunciation" criterion.	4.00	0.45
The test simulation feature helps me become familiar with the pressure of speaking tests.	4.18	0.44
The feedback provided by SpeakEasy AI helps me better understand my strengths and weaknesses	4.3	0.51
MOTIVATION OF USING SPEAKEASY AI APPLICATION	**MEAN**	**SD**
SpeakEasy AI is easy to use for speaking practice	4.13	0.51

The conversation feature motivates me to practice speaking more frequently.	4.13	0.46
After using the app, I experience reduced anxiety and pressure when speaking.	4.1	0.44
I feel more confident in communicating in English after practicing with the app.	4.03	0.48
I am satisfied with using the SpeakEasy AI app for speaking practice.	4.03	0.48
I will continue using the app for speaking practice in the future.	4.15	0.42
AVERAGE	4.05	0.29

Effectiveness of SpeakEasy AI for speaking practice

Students had positive perspectives towards the effectiveness of SpeakEasy AI in improving their speaking performance and supporting their speaking practice. The mean scores of most items in the effectiveness section was relatively high, ranging from 3.83 to 4.18 on a 5-point Likert scale.

With regard to fluency improvement, students perceived that thanks to using the application, they could speak more fluently (Mean = 4.08, SD = 0.35), retrieve vocabulary and grammar more quickly while speaking (Mean = 3.9, SD = 0.44), and reduce the frequency of pauses and hesitations during speech (Mean = 3.93, SD = 0.41). In addition, the duration of pauses and the frequency of self-corrections were reported to decline, with both items receiving positive ratings (Mean = 4.0 for the former and Mean = 4.03 for the latter).

Students also reported that SpeakEasy AI helped them employ connecting words and discourse markers better (Mean = 4.13, SD = 0.46), allowing them to speak more coherently and cohesively. Furthermore, speaking without pausing mid-thought received a similarly high score (Mean = 4.00, SD = 0.39), indicating that students were able to maintain fluid communication while performing speaking tasks to some extent.

In addition to their perceived increase in fluency, students agreed that SpeakEasy AI positively impacted their Lexical Resource (Mean = 4.03, SD = 0.53), Grammatical Range and Accuracy (Mean = 3.83, SD = 0.44), and Pronunciation (Mean = 4.00, SD = 0.45). As can be seen, grammatical range and accuracy had a relatively lower mean score compared to the other categories, demonstrating a need for further lecturer intervention in this area.

Furthermore, students found the test simulation useful because it prepared them for the anxiety caused by speaking tests (Mean = 4.18, SD = 0.44). The highest-rated item in this category was that the app benefits students by allowing them to identify their strengths and weaknesses (Mean = 4.3, SD = 0.51). These results revealed the significance of instantaneous and individualized feedback on learners' speaking awareness and self-assessment.

Motivation and Satisfaction

The majority of the items in the motivational section also had high mean scores between 4.03 and 4.15, suggesting students were highly motivated to use SpeakEasy AI to practice speaking.

Students found SpeakEasy AI to be relatively easy to use for speaking practice (Mean = 4.13, SD = 0.51). Students also reported that the conversation feature motivated them to practice speaking more frequently (Mean = 4.13, SD = 0.46). Learners also felt less anxiety and pressure when speaking after using the application (Mean = 4.1, SD = 0.44). These findings

suggest speaking on the SpeakEasy AI platform allowed learners to feel less stress when speaking.

In addition, students felt more confident speaking in English after practicing speaking on SpeakEasy AI (Mean = 4.03, SD = 0.48). Students were also satisfied with SpeakEasy AI (Mean = 4.03, SD = 0.48) and reported they would want to use this application again to practice speaking in the future (Mean = 4.15, SD = 0.42), making this item have the highest mean in the motivation section.

Motivation and Satisfaction

The majority of the items in the motivational section also had high mean scores between 4.03 and 4.15, suggesting students were highly motivated to use SpeakEasy AI to practice speaking.

Students found SpeakEasy AI to be relatively easy to use for speaking practice (Mean = 4.13, SD = 0.51). Students also reported that the conversation feature motivated them to practice speaking more frequently (Mean = 4.13, SD = 0.46). Learners also felt less anxiety and pressure when speaking after using the application (Mean = 4.1, SD = 0.44). These findings suggest speaking on the SpeakEasy AI platform allowed learners to feel less stress when speaking.

In addition, students felt more confident speaking in English after practicing speaking on SpeakEasy AI (Mean = 4.03, SD = 0.48). Students were also satisfied with SpeakEasy AI (Mean = 4.03, SD = 0.48) and reported they would want to use this application again to practice speaking in the future (Mean = 4.15, SD = 0.42), making this item have the highest mean in the motivation section.

4.3 Interview

The researcher carried out a structured interview with five students, coded as S1 to S5, who frequently used the SpeakEasy AI platform. The interviews were conducted online via Microsoft Teams. Students' experiences and perspectives on using SpeakEasy AI to practice their speaking skills were explored through their responses to four interview questions.

The first interview question explored whether students felt less anxious when practicing speaking with SpeakEasy AI and examined the underlying reasons. Overall, all interviewees reported that the application not only helped them feel less nervous and stressful but also made them more comfortable and confident during speaking practice. S1 mentioned that the app interface is user-friendly for beginners as users can choose different difficulty levels and select different AI voices. She added "*The app's voice is friendly and creates less pressure for me when speaking. It also provides specific solutions for improvement, such as correcting vocabulary, grammar, and pronunciation mistakes*". For S3, she considered the app as her friend with whom she could easily communicate and had conversations. She added: "*the app can understand what I mean even when I cannot fully express my ideas in complete sentences*". For S5, he highly appreciated the diversity of AI characters, each of whom has a different voice and different ways of asking questions and listening, which makes the app more interesting. He also emphasized an insightful observation: "*The app knows how to guide the conversation with follow-up questions and provides me with abundant vocabulary, so I do not worry about running out of ideas while speaking*"

The second question investigated whether SpeakEasy AI increased students' motivation to practice speaking English. All respondents expressed positive attitudes toward the application and agreed that it encouraged them to practice more frequently. They are all

pleased that the app is free of charge, does not limit the number of speaking attempts per day like some other applications, and allows them to practice speaking anytime and anywhere as it's available on many electronic devices. S2 stated that her favorite feature is the real-time pronunciation feedback: "*It provides immediate and detailed feedback on word stress and pronunciation. Instead of simply evaluating answers as "correct" or "incorrect," the app points out exactly which syllables need adjustment*". For S3, she was engaged with the natural conversation with the app, emphasizing the variety of topics that the app covered, helping her practice speaking more confidently without feeling bored. Another feature contributing to the increased motivation was mentioned by S4. He valued the speaking evaluation based on the four IELTS criteria: "*After the speaking session, I know my estimated band scores, recognize my weaknesses for improvement, and continue developing my strengths*".

The third question examined which speaking criterion improved the most after using SpeakEasy AI. Although the participants acknowledged their enhancement in terms of pronunciation, grammar, and vocabulary, most interviewees agreed that Fluency and Coherence was the area showing the greatest improvement. S4 reported that thanks to repeated speaking practice with the AI, he was able to remember many vocabulary and retrieve them more quickly during the speech, thereby speaking more smoothly without long pauses. In addition, according to S5, the ability of suggesting appropriate vocabulary and grammar structures enabled him to reduce the frequency of pauses and hesitations which were attributed to a lack of vocabulary and structures.

The final interview question focused on the features of the app that needed improvement and the further expectations the students had for the app. Although students generally held positive views of the application, several limitations were identified. All participants mentioned that the app sometimes interrupted them and moved on to the next question whenever they paused to think, as the system assumed that they had finished their responses. They suggest the app should enhance its speech recognition feature so that it can better determine whether the speaker has completed their response before asking another question. S1 also recommended that the pause time for speakers should be longer to allow users a few seconds to think or pause before continuing. Additionally, S2 pointed out that the app failed to recognize students' speech accurately when the pronunciation was unclear or when there was background noise. In terms of expectations, S2 stated "*I think the app should improve its pronunciation correction feature by providing phonetic transcriptions in the final feedback after each conversation*". S3 also recommended that "*The app should provide a list of vocabulary and sentence structures related to the conversation that has just been completed, so that learners can receive useful language input after finishing their speaking practice*." Moreover, S4 expected that the speaking topics should be updated and more diverse to prepare learners for IELTS test. Finally, S5 also proposed an interesting and worthwhile suggestion "*I think the app could also be suitable for secondary and high school students. It would be even better if there were chatbot characters designed to understand children's psychology and communication style, as this would make the app highly appropriate for children aged 6–10 who are beginning to develop listening and speaking skills in a new language*".

5. Discussion

After 12 weeks of integrating the SpeakEasy AI application in speaking practice activities, the findings suggested that the performance of the experimental group surpassed that of the control group in the post-test. In addition, the post-test mean scores of the EG were substantially higher than those of the pre-test, indicating that SpeakEasy AI exhibited positive contribution to students' speaking performance, particularly in terms of fluency and coherence.

These findings support previous studies on the effectiveness of AI-powered speaking applications in improving learners' oral communication skills (Kang, 2022; Qiao & Zhao, 2023; Azizimajd, 2023).

The second research question investigated students' perceptions of using SpeakEasy AI for speaking practice. The survey findings demonstrated that students were generally in favor of the application. To be more specific, more than 90% of survey participants selected "Strongly Agree" or "Agree" for most questionnaire items related to fluency development and motivation. Regarding speaking fluency, 97.5% of students agreed that SpeakEasy AI helped them speak more fluently. A closer look into the components of speaking fluency revealed that 85% of the participants reported being able to retrieve vocabulary and grammatical structures more quickly during their speech while 87.5% could reduce the frequency of pauses and hesitations as well as shortening the duration of pauses. These findings strongly resonate with Segalowitz's (2010) framework of cognitive fluency and utterance fluency, which emphasizes the importance of rapid vocabulary retrieval, reduced frequency and length of pauses in fluent speech production. Additionally, the findings confirmed the previous findings on faster retrieval of vocabulary and grammar (Qiao and Zhao, 2023), and on the ability of producing longer utterances with fewer hesitations and pauses (Azizimajd, 2023). These quantitative results were also reinforced by the interview findings, where participants reported that thanks to the repeated speaking practice, the suggestions of useful vocabulary and grammar, and the guidance of follow-up questions, SpeakEasy AI not only supported cognitive fluency through faster retrieval of language resources but also improved learners' perceived fluency by maintaining conversational flow and reducing communication breakdowns.

Apart from speaking fluency, the other speaking components also showed improvement. Nearly 90% of respondents agreed that they could develop their vocabulary and refine their pronunciation skill thanks to speaking with the app. However, fewer students achieved the comparable improvement in grammar (80%). This finding suggests that grammar development may require more explicit instruction from lecturers and guided practice rather than relying on AI-supported conversational practice.

It is also notable that the application helped familiarize students with the pressure of the speaking test. In addition, the app's feedback based on IELTS-oriented assessment features enabled them to recognize their strengths and weaknesses more effectively, as reflected in the high agreement rate of 97.5%. This finding indicated that the utilization of SpeakEasy AI app could tackle the limitations of current apps in the market which lack the standardized test-based assessment functions, such as IELTS or V-step, thereby providing learners with an all-in-one solution.

The findings also revealed high appreciation towards the app's role in reducing learners' anxiety. 95% of students agreed that SpeakEasy AI helped lower anxiety and pressure when speaking English and also expressed greater confidence in speaking English after practicing with the app. These findings were consistent with the results from earlier studies by Kang (2022) and Azizimajd (2023). Furthermore, according to Brown (2017), fluency development is more likely to occur in low-pressure and supportive learning environments where learners are encouraged to focus on meaning rather than form. This idea was prevalent in interview responses. Interviewees' answers indicated that the app's features, such as the different voice options, the friendly voice, and easy-to-use app interface created an encouraging environment for learning. Students described the app itself as "a friend" who always understood what they meant even when they answered incompletely and could offer helpful ideas, thus prompting students to try harder to speak.

Regarding motivation and satisfaction, students expressed strong positive attitudes toward SpeakEasy AI, citing easy-to-use functions and convenience for speaking practice. The overwhelming majority of students indicated they would like to use the app again in the future (97.5%), which was one of the highest-rated items on the questionnaire. These results aligned with past studies that identified affordances of AI speaking apps such as providing learners more opportunities to learn conversationally and motivating learners to practice at their own pace (Celik et al., 2025; Kang, 2022). The interview results further explained the previous findings. Students appreciated that the app was free of charge, usable on different devices, and allowed for multiple attempts per day. Students also enjoyed being exposed to various conversation topics. These qualities seemed to positively influence students' repeated speaking practice.

Along with the positive findings mentioned earlier, interviews revealed certain limitations SpeakEasy AI currently faces. Interviewees shared that the application had the tendency to cut students off from answering questions and skip to the next question whenever they stopped talking to think for a few seconds. Participants believed that SpeakEasy's speech recognition software should be able to recognize when a user is merely thinking versus when they are done with their answer. Difficulty with speech recognition was also reported when the student's pronunciation was unclear or there was background noise present. For future updates, participants wanted to see more advanced features such as phonetic spelling for pronunciation corrections, vocabulary recap after conversations, and more current IELTS speaking topics. These findings partially aligned with Celik et al. (2025), who argued that technology-based speaking applications lack certain characteristics only a human teacher can provide and cannot replace the need for a lecturer in language acquisition. While SpeakEasy AI was effective at supplementing students' speaking practice, teacher-led instruction and pedagogical assistance are still necessary for optimal learning.

6. Conclusion

This research aimed to investigate how SpeakEasy AI application influences EFL learners' speaking fluency among first-year English-major students at HUFLIT and also explore their attitudes towards using the application. Quantitative results based on post-test scores demonstrated that after 12 weeks of implementation, the experimental group scored higher than the control group. It can be concluded that using the SpeakEasy AI application as a learning tool has positive effects on students' speaking performance in terms of Fluency and Coherence. According to the survey results, practicing speaking through SpeakEasy AI application allows learners to access words and structures faster, make fewer pauses and shorten their length, and make fewer hesitations and maintain smoother communication flow.

The survey and interview results also revealed that students had overwhelmingly positive attitudes towards speaking practice with SpeakEasy AI. They agreed or strongly agreed that the application made them speak more fluently, feel less nervous, more confident, and more motivated to do speaking practice more often. They particularly enjoyed the application's instant and individualized feedback, its focus on IELTS-task-based speaking practice, conversation practice without the pressure of human judgment, and supplementary speaking practice features such as multiple voice options, vocabulary and grammar recommendations, and follow-up questions that allowed the conversation to flow naturally. These characteristics helped to reduce the learner's affective filter and promote learner engagement when practicing speaking, allowing them to feel more comfortable and confident when communicating in English. It is also noteworthy that out of the categories of speaking that saw improvement, grammar scored the lowest. This indicates that grammar may still need

explicit instruction from the teacher and more practice. Lastly, the interviews provided insight into the learners' experience with SpeakEasy AI. Students found the application flexible and free, and they appreciated that it allowed them to talk about a variety of topics that they enjoyed. Additionally, they felt that the AI characters were nice and friendly.

There are several limitations that should be addressed in this study. Firstly, the sample size was relatively small and consisted of first-year English-major students at HUFLIT only. Future research may expand the sample to students of different backgrounds and in various contexts. Secondly, the length of the intervention was only 12 weeks. As such, future research may implement longer periods of using SpeakEasy AI to examine the long-term impact of the app on learners' speaking fluency. Lastly, this study primarily focused on how SpeakEasy AI could affect speaking fluency and coherence. Studies can be done to further explore how the application can help improve other aspects of speaking, such as pronunciation, grammatical accuracy, or vocabulary building. In addition, some limitations of the application were also discovered throughout the study. Students experienced some problems with the app including occasional interruptions during speaking, speech recognition inaccuracies, outdated IELTS speaking topics, and a need of post-conversation phonetic transcription and vocabulary lists. Therefore, the findings suggest some practical implications for the app developers. Firstly, the speech recognition system should be enhanced to distinguish between thinking pauses and completed responses. Rather than proceeding to the next question, the app should allow the users to customize their response time depending on their proficiency levels. Secondly, the app should improve speech recognition accuracy especially with non-native accents and in environments with background noises. When it comes to pedagogical implications, while waiting for the developers to improve these features, some practical recommendations can be offered to educators. With regard to the app's unwanted interruption, students should divide their longer responses into shorter ones. If the app interrups their answer, they should ask the app to continue the previous conversation and complete the rest of the answer. Regarding the speech recognition issue, students should speak in a quiet environment, check their microphone settings before practice, try to articulate their words clearly, and avoid using the app late at night during peak hours. In addition, the teacher should integrate AI-generated feedback into classroom instruction by asking students to review AI feedback reports and identify repeated pronunciation and grammar mistakes so that the teacher can provide further instruction to help students improve their weaknesses. Furthermore, teachers should provide students with IELTS updated speaking topics and topic-related vocabulary, and encourage students to practice these topics with the app.

Overall, the findings of this study contribute to the growing body of research supporting the integration of AI-powered applications into EFL speaking instruction. SpeakEasy AI application is proved to be an effective supplementary tool that combines speaking practice, test simulation, and criteria-based feedback in a single platform. Therefore, SpeakEasy AI can be considered a promising approach for enhancing speaking fluency and promoting learner motivation in EFL contexts.

References

Alaraj, M. (2017). EFL speaking acquisition: Identifying problems, suggesting learning strategies and examining their effect on students' speaking fluency. *International*

Journal of Social Sciences and Humanities Invention, 4(1), 3215–3221. https://doi.org/10.18535/ijsshi/v4i1.05

Amaoh, S., & Yeboah, J. (2021). The speaking difficulties of Chinese EFL learners and their motivation towards speaking the English language. *Journal of Language and Linguistic Studies, 17*(1), 56–69. Doi: 10.52462/jlls.4.

Azizimajd, H. (2023). Investigating the impacts of voice-based student-chatbot interactions in the classroom on EFL learners' oral fluency and foreign language speaking anxiety. *Technology Assisted Language Education, 1*(2), 61–83. https://tale.razi.ac.ir/article_2732_fc8d876a674e0da4ec799ca44899b5b1.pdf

Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning, 50*(2), 311–341. https://doi.org/10.1111/0023-8333.00119

Bárkányi, Z. (2021). Motivation, self-efficacy beliefs and speaking anxiety in language MOOCs. *ReCALL, 33*(2), 143–160. https://doi.org/10.1017/S0958344021000033

Brown, P. S. (2017). What is fluency and how do we develop it? *TESL Ontario. CONTACT Magazine, Early View (Nov)*, 5–7. https://contact.teslontario.org/wp-content/uploads/2017/11/03Brown-Fluency.pdf

Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics, 8*(1), Article 102418. https://doi.org/10.29140/ajal.v8n1.102418

Engeström, Y. (1987/2015). *Learning by expanding: An activity-theoretical approach to developmental research*. Cambridge University Press. https://doi.org/10.1017/CBO9781139814744

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning, 37*(1–2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

Kang, H. (2022). Effects of artificial intelligence (AI) and native speaker interlocutors on EFL learners' speaking ability and affective aspects. *Multimedia-Assisted Language Learning, 25*(2), 9–43. doi:10. 15702/mall.2022.25.2.09

Karim, S. A., Hamzah, A. Q. S., Maudyna, N., Prianti, J., & Sihole, I. G. (2023). Promoting EFL students' speaking performance through ELSA Speak: An artificial intelligence in English language learning. *Journal of Languages and Language Teaching, 11*(4), 655–668. https://e-journal.undikma.ac.id/index.php/jollt/article/view/8958/4932

Kumalasari, C., & Wulandari, F. (2025). The perception of using ELSA Speak application to

improve speaking skills of informatics students. *Journal of English Language Teaching and English Linguistics, 10*(2), 23–33. https://doi.org/10.31316/eltics.v10i2.6201

Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis (TTS): Focus on French liaison. *Computer Assisted Language Learning*, *30*(3-4), 325-342. DOI: 10.1080/09588221.2017.1312463

Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL, 31*(3), 293–311. https://doi.org/10.1017/S0958344019000065

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System, 57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

Pabro-Maquidato, I. M. (2021). The experience of English speaking anxiety and coping strategies: A transcendental phenomenological study. *International Journal of TESOL Studies, 1*(2), 45–64. https://i-jte.org/index.php/journal/article/view/32

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology, 14*, Article 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge. https://doi.org/10.4324/9780203851357

Tavakoli, P., Nakatsuhara, F., & Hunter, A.-M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *Modern Language Journal, 104*(1), 169–191. https://doi.org/10.1111/modl.12620

Bionote

Nguyen Thi Huyen Tran is currently an EFL lecturer in the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). She holds a Master's degree in Applied Linguistics from Curtin University, Australia. She has over 14 years of experience teaching IELTS to non-English-major students at Ho Chi Minh City University of Technology - Vietnam National University Ho Chi Minh City (HCMUT-VNU). Her research interests include Teaching English to Speakers of Other Languages (TESOL) and AI-assisted language learning.

Email: trannth1@huflit.edu.vn

Acknowledgement

The author would like to express her sincere gratitude to Mr. Nguyen Le Nhu Anh, an English lecturer at Ho Chi Minh City University of Education and the founder of Edtechcorner, a platform that provides free AI-powered applications to support English language teaching and learning for both teachers and students. The SpeakEasy AI application investigated in this study was also developed on this platform.

Special thanks are also extended to Mr. Nguyen Le Nhu Anh for his enthusiastic support and for providing valuable background information about SpeakEasy AI, which greatly

contributed to the development of the theoretical framework and the successful completion of this research.

#149 A FEASIBILITY STUDY OF A GLOBALIZATION AND INTERCULTURAL COMMUNICATION PROGRAM AT THE UNIVERSITY OF LANGUAGES AND INTERNATIONAL STUDIES, VIETNAM NATIONAL UNIVERSITY HANOI: A QUINTUPLE HELIX MODEL

Tran Thanh Nhan[1], Vu Thi Kim Ngan[2], Chu Thanh Van[3]

Abstract

This feasibility study on launching the Globalization and Intercultural Communication (GIC) program was initiated to advance the International Relations majors in a Vietnamese transformative context. The study was conducted using a mixed-methods approach, guided by Bronfenbrenner's (1979) multi-layered quintuple-helix ecological system and a tri-dimensional needs analysis model (incorporating Target Situation, Present Situation, and Critical Needs Analysis). It involves documentary analysis of official directives and four existing curricula together with a large-scale needs assessment ($N = 1562$) encompassing five stakeholder groups. Three major findings were identified: a macro promotion of globalization and internationalization in education in governmental policies; a meso institutional readiness supported by highly compatible ongoing courses, sufficient human resources, and the

[1]*Lecturer, Faculty of English Language and Culture, VNU, University of Languages and International Studies, Hanoi, Vietnam; Email: thanhantn81@ulis.vnu.edu.vn*
[2]*Officer, Office of Research and Partnerships, VNU, University of Languages and International Studies, Hanoi, Vietnam*
[3]*Lecturer, Faculty of English Language and Culture, VNU, University of Languages and International Studies, Hanoi, Vietnam*

university ecosystem; and a major interest at the exo and micro layer, suggesting a paradigm shift from language mastery to intercultural competence. GIC curriculum would be competence-based, developing soft, thinking, and professional skills, intercultural competence, and digital/AI literacy. Moreover, it should orient towards practicality and applicability, offering corporate and professional training, fostering multilingualism and intercultural communication, and adopting multidisciplinary or hybrid major–minor structures.

Keywords**:** internationalization, needs assessment, curriculum development, multi-layered ecosystem, a quintuple helix model

1. Introduction

Amid shifting geopolitical landscapes and swift socio-technological transformation, higher education institutions are confronting a dual challenge of preserving their identity while pursuing internationalization (de Wit, 2020; OECD, 1999; Tight, 2022). To address this obstacle, they must renew themselves by upgrading their study programs to address potential opportunities and threats. This study, therefore, examines the feasibility of introducing a novel study program: Globalization and Intercultural Communication (GIC) at the University of Languages and International Studies, Vietnam National University, Hanoi (VNU-ULIS).

Founded in 1955, VNU-ULIS has historically evolved from a specialized language training academy into a comprehensive center for language teacher training, applied linguistics, and international studies. Since its 1993 membership in Vietnam National University, the university has been rebranded as the University of Languages and International Studies to highlight its mission in the internationalization process. VNU-ULIS students and faculty have been appointed to represent the Vietnamese youth and academics to greet numerous delegations of high-ranking leaders. The fact that many alumni hold important positions in the government, the Ministry of Foreign Affairs, the Vietnamese Embassy, leading international universities, and multinational corporations is living evidence of the university's international integration and development.

The institutional drive to launch this new GIC program coincides with a major milestone: the VNU-ULIS's 70th anniversary and the structural consolidation of its international studies. The increasingly significant role of Vietnam in the global arena across socioeconomic, cultural, and political fields is the motive behind this transformative move to prepare agentic human resources for international job opportunities and responsibilities. This research, therefore, was conducted to examine the facilitating conditions created by the Vietnamese government and the university to promote the new program and capture various stakeholders' aspirations via an extensive needs assessment. These objectives are to be addressed using the quintuple helix model that encapsulates the dynamic partnership among five "houses": the state, the university, enterprises, researchers, and the public. To evaluate the overarching feasibility, three research questions are formulated as follows:

(1) What governmental policies and initiatives support the GIC training program?

(2) How can VNU-ULIS address the requirements for launching the new program in terms of curriculum, human resources, and facilities?

(3) What are the stakeholders' interests and anticipation of employment opportunities for GIC graduates?

(4) What recommendations are proposed for GIC curriculum development?

2. Literature review

Feasibility study

Feasibility studies, or pilot studies, are often used to determine whether a new idea, intervention, program, process, or measure is suitable for further testing. In other words, this method allows researchers to assess whether the idea or outcome is appropriate and sustainable (Bowen et al., 2009, p. 2). These studies are often conducted before a main study or in the initial stages of large projects to estimate key parameters necessary for the development of the main research (Arain et al., 2010). They can focus on various aspects such as acceptance, needs, implementation, applicability, adaptation, integration, expansion, or limited-efficacy testing (ibid.). The goal is to answer three main questions: *Can it work? Does it work? Will it work?* (Bowen et al., 2009; Orsmond & Cohn, 2015).

However, this study is set to answer two questions: *Can it work?* and (*How) will it work?* The former is addressed by an analysis of the State and Government policies on globalization and educational international integration goals and the VNU-ULIS capacity to implement the program; the latter is addressed by career-oriented needs as perceived by various stakeholders. Assumingly, there can be consistency among the state's leadership, the university's capacity, and stakeholders' aspirations for GIC curriculum development. Meanwhile, career-oriented needs would be translated into a system of learning units covering knowledge, skills/competencies, and qualities, which may serve as key components of the new curriculum (Agrawal et al., 2016).

The quintuple helix model

The term "triple helix model" has recently been associated with the promulgation of Resolution No. 57-NQ/TW dated December 22, 2024, of the Politburo, which emphasizes the combined efforts of the Party, the entire political system, entrepreneurs, businesses, and the citizens in science, technology, innovation, and national digital transformation. In this technological revolution, the citizens and businesses are considered the center, the subjects, the main resources, and the driving force; scientists are the key factor; and the State plays a leading role in creating the most favorable conditions (Government Electronic Newspaper, 2026). It can be claimed that this tripartite cooperation has transformed Resolution 57-NQ/TW into an endogenous driving force for national, transformative development.

In actuality, Resolution No.57-NQ/TW is being widely implemented in technology, education, and research, with groundbreaking solutions (Lan, 2025; Thach, 2022). Ho Chi Minh City National University is a pioneer in implementing this triple helix model (State - University - Enterprise) to quickly concretize the Resolution into strategies to develop high-quality human resources, enhance global competitiveness, and modernize digital technology infrastructure (Nhan Dan Online Newspaper, 2025). This model has also been implemented elsewhere to evaluate the effectiveness of AI training programs for teachers (Van et al., 2026) or develop high-quality semiconductor human resources (Le et al., 2025). At Tan Trao University, it is adopted to transform the university into a knowledge hub, connecting and spreading innovations for sustainable development (Ha & Cuc, 2025). At Hanoi University of Theatre and Film, the model is adopted to create a mechanism for ordering scientific research projects and creative scientific products to develop the cultural industry (Thi & Binh, 2026).

However, the triple helix model apparently neglects the role of citizens and scientists, as mentioned in Resolution 57-NQ/TW. Other foreign studies also illuminate more complex multi-stakeholder collaborative models in research on social innovation, entrepreneurial ecosystem, and sustainable development (Borrero & Yousafzai, 2024; Carayannis & Campbell,

2012; Cloitre et al., 2023; Morawska-Jancelewicz, 2022). Therefore, the "quintuple helix model", five-stakeholder collaboration: State - University - Researcher/ Expert - Employer - Citizen, was adopted for this study. The difference is that they are clustered into four layers of an ecosystem.

Multi-layered ecosystem conceptual framework

The multi-tiered quintuple helix model based on Bronfenbrenner's (1979) theory of human development's ecosystem lays the theoretical foundation for this study. The framework allows consideration of the reciprocal impacts among policies, institutions, enterprises, academics, and individuals on curriculum development corresponding to the macro, meso, exo, and micro levels. In this superstructure, the macro level focuses on national policies regarding globalization and internationalization of education. The meso level presents the institutional capacity as reflected in the curricula, human resources, and facilities. The exo and micro levels involve employers, industry professionals/researchers who may offer the most authentic perspectives on industry development trends, and students, graduates, and high school students, the direct and potential beneficiaries of the training programs. This multi-layered ecosystem illuminates the interconnections among government policies, institutional implementation, and the occupational requirements of the GIC program.

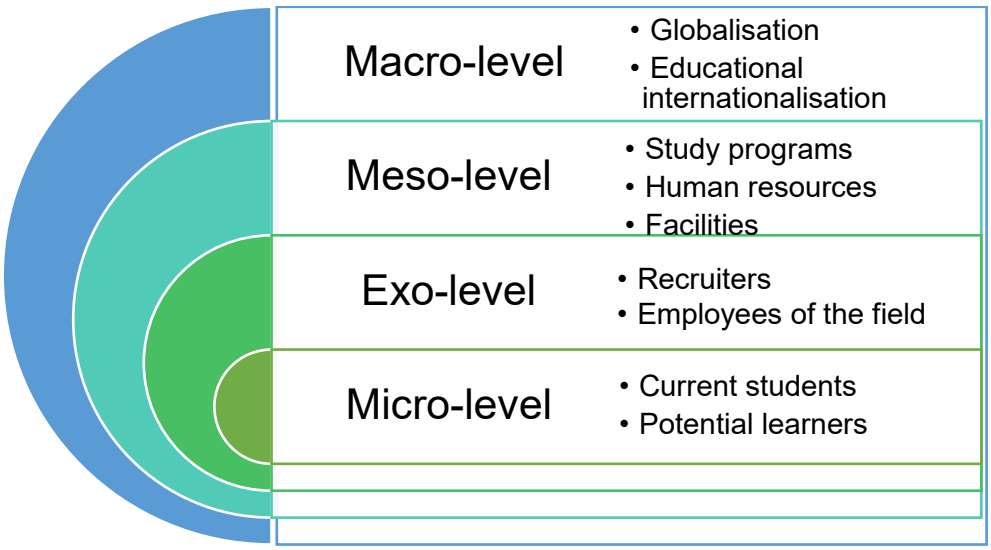


Figure 1: Multi-layered conceptual framework based on Bronfenbrenner's Ecosystem Theory (1979)

3. Research methodology

This study employs a mixed-methods approach in a convergent parallel design to corroborate qualitative and quantitative data. The study is grounded in a multi-tiered ecosystem conceptual framework (Bronfenbrenner, 1979) as mentioned above, and adopts an analytical framework of a tri-dimensional model incorporating Target Situation Analysis (TSA), Present Situation Analysis (PSA) (Chambers, 1980; Hutchinson & Waters, 1987), and Critical Needs Analysis (CNA) (Benesch, 2001).

Data collection

Qualitative data. The qualitative data pool consists of (a) official documents from the state, government, and ministries; (b) four existing curricula and internal evaluation reports on human resources and facilities; and (c) approximately 500 responses to the survey's open-

ended questions. Documentary and thematic analysis were adopted to process and report this data type.

Quantitative data. They were harnessed from a large-scale needs assessment based on a Cognito survey. The questionnaire construct was adapted from a similar prototype designed for the Transnational Cultural and Media Studies program launched by VNU-ULIS in 2023; however, this versatile model incorporated five stakeholder groups into a single survey and targets more experienced, reflective participants who can provide insights into the actual market demands for this major. The snowball sampling technique and social media platforms were used to administer the questionnaire to potential participants.

Participants. After more than 6 months, the form recorded 1580 entries; yet the data cleaning yielded a sample of 1562 valid responses. Notably, the large standard deviations, ranging from 1.30 to 1.62, indicated heterogeneity among stakeholders' opinions. Demographically, the sample included 735 high school students (47.1%), 497 university students (31.8%), 236 employees (15.1%), 40 employers (2.5%), and 54 researchers/experts (3.5%). 78.4% of participants were aged 15-22; those aged 23-40 accounted for 14%; and the group over 41 accounted for 7.6%. Females constituted 64.9%, males 29.9%, and the remaining 5.2% did not want to disclose their gender. Regarding educational level, 67.0% classified themselves as currently studying/having completed high school; 22.7% have a bachelor's degree. Holders of a Master's degree or higher accounted for 7.4%, including 79 Master's degree holders, 31 PhD holders, and 6 Professors/Associate Professors. 2.9% selected the option "Other" to clarify their educational background. Their fields of interest were ranked as follows: languages (26.3%), economics (22%), education (16.8%), technology (9.1%), politics (7.7%), and history (7.4%). Additionally, 168 (10.7%) participants expressed other interests, including graphic arts, real estate, security, diplomacy, international relations, tourism, services, engineering, and robotics.

Career-oriented needs analysis

Long (2005) contended that needs analysis, especially large-scale research, is of high significance in curriculum development to identify the gap between learners' current capabilities (present needs) and the stringent requirements of the target situation (target needs), thereby establishing a basis for developing appropriate objectives, content, and teaching methods. Numerous needs assessment studies have been conducted to develop language curricula, particularly English for Specific Purposes (Andi & Arafah, 2017; Boroujeni & Fard, 2013; Dewi et al., 2023; Kaewpet, 2009; Li, 2014; Long, 2005; Richards, 2009). However, needs surveys to develop an entire undergraduate content-based program are rarely found in the literature.

The needs assessment for the GIC curriculum was, therefore, located within a rigorous analytical framework that combines traditional and modern approaches: TSA, PSA (Chambers, 1980; Hutchinson & Waters, 1987), CNA (Benesch, 2001). Recent studies highlight the combination of TSA and PSA to identify gaps in curriculum development (Dou et al., 2023; Rahman, 2015). Meanwhile, CNA may ensure fairness and democracy by recognizing the contributions of students and working professionals as central subjects of the training process.

Present Situation Analysis (PSA) delves into the macro and meso layers of the human development ecosystem to illuminate current sociopolitical and institutional factors regarding globalization and educational integration. More specifically, PSA would gather information on the state's current political guidelines that could provide the foundation and impetus for

developing the GIC curriculum and reflect VNU-ULIS's capacity to meet the demand to open this major in terms of training programs, human resources, and facilities.

Targeted Needs Analysis (TSA) focuses on identifying what learners need to do in the target situation (Dai, 2026). This is an objective, needs-oriented, product-oriented approach that seeks to specify language skills and professional knowledge necessary for students to perform future jobs (Arnó-Macià et al., 2020; Dou et al., 2023). This career-oriented needs analysis strives to shape objective professional competencies, language skills, and operational benchmarks required of graduates in their future target workplaces.

Critical Needs Analysis (CNA) emphasizes the social, political, and power dynamics in the learning process (Benesch, 2001). This study, therefore, expands its scope to examine governmental and institutional documents and invites contributions to drafting the new curriculum from multiple stakeholders. This multi-dimensional perspective reflects a dynamic, flexible alignment among state governance, institutional operationalization, and citizens' implementation. It also empowers learners to contribute to curriculum development, which is tailored to their future career prospects.

Findings

Governmental policies and initiatives to support the GIC training program

Globalization as a policy. Vietnam joined the World Trade Organization in 2007. Since then, globalization and international cooperation have been regarded as the main driving force for the country's comprehensive development. In 2021, as the country transitioned into a new decade of development, the Communist Party of Vietnam once again affirmed in the Party's 13th National Assembly Document that international integration is an inevitable trend alongside the development of a socialist-oriented market economy (Đảng Cộng sản Việt Nam, 2021). This document outlines policies aimed at proactive and comprehensive international collaboration, enhancing the nation's position, and developing culture in line with national identity.

Also at this turning point, the Government developed the Socio-Economic Development Strategy 2021–2030, with a vision to 2045, with initiatives to innovate the growth model, integrate deeply through participation in global value chains, promote digital transformation, and develop high-quality human resources (Bộ Chính trị, 2021). In the Cultural Strategy to 2030, the Ministry of Culture, Sports and Tourism proposed ways to develop culture by promoting Vietnamese values, preserving national identity, and absorbing the best of humanity, which was approved by Decree 1909/QĐ-TTg (Thủ tướng Chính phủ, 2021) and reinforced in Decree 80 NQ/TW (Bộ Chính trị, 2026). Most recently, the Vietnam 2045 Report by the Ministry of Planning and Investment analyzed global economic trends and policy implications to adapt to geopolitical changes, leverage technological opportunities, and enhance national competitiveness (Viện Chiến lược phát triển, 2024).

Policies on international cooperation in education. In 2018, the Government implemented an open-door policy for education, attracting foreign investment into the sector and encouraging joint training programs through Decree 86/2018/NĐ-CP (Chính phủ Việt Nam, 2018). The documents of the 13th National Assembly of the Communist Party of Vietnam in 2021 affirmed that education is a top national priority and linked educational development to international integration to further promote cooperation, develop high-quality human resources, and strengthen academic exchange. Also in that year, the Ministry of Education and Training issued the Education Development Strategy 2021–2030, with a vision to 2045, demonstrating a determination to fundamentally and comprehensively reform

education, further enhance the role of international cooperation through training partnerships, attract international human resources, and develop a global research network (Báo Điện tử Chính phủ, 2025).

In 2024, Resolution 57-NQ/TW on breakthroughs in national science, technology, innovation, and digital transformation accelerated the country's development and international integration with the power of science and technology, aiming to enhance the nation's position through comprehensive development of national resources, from policies, institutions, infrastructure, and human resources to international cooperation (Ban Chấp hành Trung ương, 2024). Most recently, in Resolution 71-NQ/TW, the Political Bureau concretized the policy of developing higher education in the era of international integration through a strategy of increasing autonomy and building research universities with the goal of comprehensive integration, affirming the pioneering role of Vietnamese education among the top 20 countries in the world, with at least five universities in the Top 100 globally (Ban Chấp hành Trung ương, 2025).

Also in 2025, the Prime Minister issued Decision 2371/QĐ-TTg approving the Project "Making English the second language in schools in the period 2025–2035, with a vision to 2045". This decision has great significance for promoting international cooperation through education, enhancing the competitiveness of human resources, creating conditions for the younger generation to access global knowledge, and aligning with the national development strategy to 2045 (Thủ tướng Chính phủ, 2025). Thus, alongside the strengths of digital transformation, the state and government also affirm the indispensable role of language in the strategy of international integration and building national identity with the vision to 2045.

To recapitulate, the governmental policies signify a strong motive to establish globalization and educational integration as top developmental goals, steering the nation towards modernization, digitalization, and international collaboration. The trio of technological transformation, more extensive global outreach, and English as a second language policy, as postulated by Resolution 57-NQ/TW, Resolution 71-NQ/TW, and Decision 2371/QĐ-TTg, acts as a catalyst for a new era of development in Vietnam.

VNU-ULIS's capacity regarding curriculum, human resources, and facilities

Curriculum. Currently, the university offers training in eight of the world's most popular languages (English, Chinese, Korean, Japanese, French, German, Russian, and Arabic) in two majors: education and/or linguistics. In addition, the university recently launched two new training programs: Transnational Culture and Media Studies and Teaching Vietnamese as a Foreign Language. According to the 2025 plan, VNU-ULIS planned to open six more (Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội [ĐHNN-ĐHGQHN], 2025a). The GIC program is anticipated to intensify internationalization, aligning it with the state policies on globalization and international integration, as synthesized in RQ1.

According to the self-assessment report on the undergraduate English Language program (ĐHNN-ĐHQGHN, 2025b), the curriculum was established based on the university's core values of "innovation, national responsibility, sustainable development" (p. 23) and the educational philosophy of "social construction, comprehensive development, autonomy, liberation, dynamism, and creativity" (p. 17). It serves the missions of "contributing to the development and integration, building and protecting the country through training high-quality human resources, conducting scientific research, implementing innovation, transferring knowledge, providing products and services that meet social needs according to international standards in the fields of foreign language education, linguistics, international studies and some

related social sciences and humanities" (p. 23). The vision for 2045 reflects the university's aspiration to become a highly reputable institution in its fields, not only domestically but also regionally and globally.

International studies is currently a sub-specialization within the English Language major (ĐHNN-ĐHQGHN, 2017). This orientation is clearly stated in Objective 2 as a crucial interdisciplinary knowledge base for creative professional practice and is also manifested in multiple program learning outcomes. However, a more extensive documentary analysis of the four current curricula - English Language, English Pedagogy, Transnational Culture and Media Studies, and Teaching Vietnamese as a Second Language (ĐHNN-ĐHQGHN, 2017, 2019, 2023, 2025) revealed that 92 out of 162 studied subjects (56,8%) (equivalent to 276 credits) have a high degree of compatibility with the GIC orientation. This suggests that VNU-ULIS's training programs are mostly internationalized and can cater to the needs for international integration (Table ***26***).

Table 26.

Level of Compatibility with the GIC Orientation of Subjects in the Four Current Curricula

Block of knowledge	Number of subjects	Level of compatibility		
		NA/1	2	3
General knowledge	***41***	***2***	***14***	***25***
Vietnamese studies	13	2	1	10
Social-emotional competence	10	0	5	5
Digital & multimodal competence	9	0	2	7
Research competence	9	0	6	3
Linguistic and Cultural knowledge	**37**	**0**	**2**	**35**
Linguistic competence	23	0	2	21
International studies	14	0	0	14
Specialized knowledge	**84**	**8**	**44**	**32**
Education	16	6	10	0
Teaching Vietnamese as a foreign language	10	0	10	0
Translation-Interpretation	9	0	1	8
Linguistics-Culture	16	1	12	3
Administration	6	0	2	4
Tourism	8	0	6	2
Transnational Cultural and Media Studies	8	1	3	4
Global Development	11	0	0	11
Total	**162**	**10**	**60**	**92**

N/A: Not available

1. Low compatibility
2. Average compatibility level

3. High level of compatibility

However, courses aligned with the GIC orientation are spread across various knowledge blocks. 60 out of 92 highly compatible courses (65.2%) are within the General and Linguistic and Cultural knowledge blocks, mostly in Linguistic competence (21 subjects) and the International Studies strand (the English Language program) (14 subjects). The remaining (32 subjects) belong to the specialized knowledge block, largely concentrating on the Global Development strand (Transnational Culture and Media Studies program) (11 subjects) and the Translation-Interpretation strand (the English Language program) (8 subjects). Among those highly compatible, specialized subjects, only four are classified as highly practical: *Sustainable Development, Culture and Politics in Business, International Business, and International Commercial Law*; the remaining are highly theoretical and academic. The findings suggest consolidating GIC-oriented subjects into a single study program to strengthen this major.

Human resources. According to ĐHNN-ĐHQGHN (2025), VNU-ULIS's faculty possesses admirable expertise in foreign language education, linguistics, international studies, and related disciplines. Over 70% are doctoral or master's degree holders, many of whom were qualified by prestigious universities in Europe, the United States, Japan, South Korea, China, and Australia. With diverse academic backgrounds and international experience, they work as professional trainers and continuously retrain themselves to stay ahead of the times. Moreover, many faculty members have participated in collaborative projects, academic conferences, and joint research programs to advance their research skills. Their works have been published in prestigious journals, focusing on foreign languages, international studies, regional studies, and other interdisciplinary fields. Faculty members also engage in consulting activities and collaborate with businesses, international organizations, diplomatic agencies, ministries, and departments to find internship and employment opportunities for their students. Their sense of responsibility, dynamism, and integration is crucial in affirming VNU-ULIS's position as a leading training and research institution in the fields of foreign languages and international studies.

In addition to its six permanent PhD and 28 Master's degree holders in international relations, literature and culture, development economics, education, and communication who can serve as GIC course leaders and instructors, VNU-ULIS can also mobilizes a force of visiting lecturers (11 members) including associate professors and PhDs in the fields of politics, diplomacy, linguistics and script studies (Chinese), Oriental studies, philosophy, international economics, international business, financial economics, and business administration. This combination ensures the program's multidisciplinary nature and the connection between theory and practice. This is an important foundation for the development of the GIC program. Furthermore, the GIC program's implementation and management require the participation of program administrators, student management specialists, academic advisors, and technical staff. The university's functional departments and offices manage and support students with administrative tasks, academic advisors offer personalized study strategies, and the technical staff ensures the smooth operation of technological systems for learning, research, and assessment. They are professionally trained to ensure standardized official student services.

Over more than 70 years of formation and development, the university has built a dedicated faculty with sound professional credentials, varied practical experience, and resilient research capabilities. These generations of faculty members have contributed significantly to training high-quality human resources, while simultaneously building their academic prestige. Besides, the visiting lecturers, administrative and student support staff, and international collaboration also play a crucial role in implementing the GIC curriculum. The combination of

in-depth knowledge, practical experience, quality service, extensive cooperation, and a spirit of innovation will ensure the program's quality and prestige, while strengthening the trust of students, society, and international partners.

Facilities. According to ĐHNN-ĐHQGHN(2022), VNU-ULIS is well-equipped with a modern lecture hall system, a multilingual library, smart classrooms, and an information technology system that supports standard, online, and blended learning. According to internal reports, the school operates seven main building blocks serving teaching and learning activities, with a diverse range of classrooms, from standard classrooms of 30-50 m² (170 rooms) to auditoriums of 75-120 m² (22 rooms). All classrooms are equipped with air conditioning, energy-efficient lighting, and audiovisual equipment suitable for teaching and learning languages in a highly interactive environment. Additionally, the school has several high-tech functional rooms: an audiovisual practice room, 2 online learning rooms, 17 language laboratories, and a large lecture hall. These diverse and regularly upgraded classrooms and lecture halls can effectively meet modern teaching requirements, for example, multimedia presentations, seminars, workshops, and simulated United Nations conferences.

In terms of digital facilities, the school also has smart classrooms, labs, online learning rooms, a learning materials production studio, a multilingual library, and a modern IT system. The entire school currently owns over 660 computers used for administrative and educational purposes, 169 projectors, and 705 air conditioners. Wi-Fi is available in all work areas and lecture halls, offering faculty, researchers, and students free internet access. The university's internal network connects to the VNU network and three external systems (FPT, VNPT, and Viettel). To effectively support administrative and management activities, the university operates an official website (www.ulis.vnu.edu.vn) upgraded and reorganized to be more of a management than an information network, an e-office document management software system (https://eoffice.vnu.edu.vn/) connected seamlessly with the VNU network, and a training and student affairs management network (www.daotao.vnu.edu.vn) allowing the university to create timetables, manage records, calculate grades, classify students, and award scholarships.

Furthermore, an online learning management system (LMS) to support the blended learning model has been implemented for many subjects, reducing the burden on physical facilities. The university has also actively invested in technological solutions such as Zoom meetings, Microsoft Teams, Turnitin, and ULIS chatbot to optimize the service quality for learners. The library and the IT & Learning Resources Department retain over 29,000 documents and are connected to national and international databases to ensure rich and up-to-date learning resources. Students and lecturers have access to the physical reading room in the French Faculty building, as well as the in-person and online reading rooms at the VNU-LIC Library and Digital Knowledge Centre. These digital solutions have significantly improved the functionality of classrooms, lecture halls, and information technology infrastructure managed by the Resource Management Centre.

The VNU-ULIS ecosystem also features administrative areas (A1, A5), shared workspaces for lecturers and students (Homies), the International Conference Centre, a physical training and practice area, dormitories, and open spaces to create the best campus experience for students. The administrative buildings A1 and A5 house the offices for faculty leaders and specialized departments. The "Homies" system creates a modern and creative self-study environment for students and lecturers. The International Conference Centre hosts conferences, seminars, and workshops. The dormitories provide sufficient living space and accommodation for resident students. The open spaces offer suitable amenities. The physical training and practice area provides facilities and grounds for cultural, artistic, and sports

activities. The university has recently invested in installing an additional elevator in the main lecture building A2 and opening new access points in lecture building B3 to facilitate faster transitions.

To recapitulate, at the meso layer, VNU-ULIS has sufficient capacity and resources to launch a new GIC program. The four existing curricula contain 92/162 highly compatible subjects. The human resources team, comprising 45 teaching members and other student support staff, can ensure the program's smooth operation. The university is also constantly improving both tangible and intangible infrastructure. Therefore, VNU-ULIS has all the necessary conditions in terms of curriculum, human resources, and facilities to organize and successfully and sustainably implement the GIC program.

The stakeholders' interests and anticipation of employment opportunities for the GIC graduates

Interests in the GIC training program. The International Relations major and the GIC program were recognized by 67% of participants, indicating the brand name's relevance in Vietnam's globalized society. This is considered a favorable starting point to launch the new program and *build* brand recognition. 57% of participants showed interest in the GIC training program, of which 21% expressed strong interest. Notably, the number of neutral (31%) and moderately interested responses (36%) was quite high, indicating the program's potential to attract even more attention with stronger promotion.

To be able to pursue this program, participants identified foreign language skills (M = 3.68, SD = 1.62), understanding of the language and culture associated with that language (M = 3.67, SD = 1.49), and communication knowledge and skills (M = 3.53, SD = 1.37) as the most important knowledge/skills to develop. The second most important group includes world civilization history (M = 3.48, SD = 1.31), understanding of globalization (M = 3.46, SD = 1.43), and intercultural communication skills (M = 3.45, SD = 1.48). Geopolitics (M = 3.42, SD = 1.36), political economy (M = 3.32, SD = 1.37), and international event analysis skills (M = 3.32, SD = 1.38) are the subjects with the lowest mean scores, but this would be associated with the increasingly difficult level of specialized knowledge and skills rather than their declining significance. Apart from the knowledge/skill groups mentioned, learners also need to perfect others necessary for their future work (M = 3.42, SD = 1.39) (Table 27).

Table 27

Knowledge/Skills that Need to be Developed

Knowledge/Skills that need to be developed	*M*	*SD*
Foreign language skills	3.68	1.62
Understanding the language and the culture of that language.	3.67	1.49
Communication knowledge and skills	3.53	1.37
Understanding globalization	3.46	1.43
Cross-cultural communication skills	3.45	1.48
History of world civilization	3.48	1.31
International event analysis skills	3.32	1.38
Geopolitics	3.42	1.36
Political economy	3.32	1.37
Other	3.42	1.39

Career opportunities for GIC graduates. 64% of respondents believed the GIC program could increase their career prospects, with 27% expecting opportunities to rise by 2-19%, 25% predicting a 20-30% increase, and 12% optimistic about a 31-50% surge. On the negative side, 21% thought career opportunities might decline, and 15% had no clear answer to this question.

According to the closed-ended question, the most suitable career paths for GIC graduates are language, culture, and international relations research (M = 3.72, SD = 1.50), translation (M = 3.65, SD = 1.42), teaching (M = 3.64, SD = 1.41), and diplomacy (M = 3.60, SD = 1.48). The second most suitable career paths are development project implementation (M = 3.52, SD = 1.36) and communication/marketing (M = 3.50, SD = 1.32). Politics is considered the least suitable profession (M = 3.41, SD = 1.38), possibly because it is a high-status profession and is often suited to exceptional individuals (Table 28). Notably, less than 10% of survey participants did not comment on this question.

Table 28

Suitable Occupations for GIC Graduates

Suitable occupations for GIC graduates	Valid answers	Don't know	*M*	*SD*
Teaching	1438	124	3.64	1.41
Translation	1437	125	3.65	1.42
Research in language, culture, and international relations	1449	113	3.72	1.50
Implement development projects	1416	146	3.52	1.36
Communications/Marketing	1443	119	3.50	1.32
Politician	1389	173	3.41	1.38
Diplomat	1433	129	3.60	1.48

When it comes to career opportunities for GIC graduates, the participants reflected a more cautious attitude with lower mean scores than in the previous question. Accordingly, significant career opportunities can be found in language, culture, and international relations research (M = 3.36, SD = 1.55), international organizations, non-governmental organizations (NGOs) (M = 3.34, SD = 1.50), development projects (M = 3.28, SD = 1.39), diplomacy (M = 3.28, SD = 1.51), communication/marketing (M = 3.27, SD = 1.36), the foreign-invested sector (M = 3.25, SD = 1.42), and research institutes/centers/higher education institutions (M = 3.25, SD = 1.42). Occupations with less promising career prospects for GIC graduates include translation (M = 3.24, SD = 1.48), private sector (M = 3.23, SD = 1.33), public sector (M = 3.22, SD = 1.31), public administration (M = 3.21, SD = 1.30), politics (M = 3.20, SD = 1.36), and teaching (M = 3.19, SD = 1.45). The number of participants who answered "I don't know" ranged around 10% (

Table **29**).

Table 29.

Career Opportunities for GIC Graduates

Career opportunities for GIC graduates	Valid answer	Don't know		*M*	*SD*
Teaching	1449	113	9	3.15	1.4
Translation	1446	116	4	3.28	1.4
Studies in language, culture, and international relations.	1475	102	6	3.35	1.5
Implement development projects	1422	140	8	3.29	1.3
Communications/Marketing	1440	122	7	3.26	1.3
Politics	1396	166	0	3.26	1.3
Diplomatic	1438	124	8	3.21	1.5
State-owned economic sector	1376	186	2	3.21	1.3
Private sector	1391	171	3	3.23	1.3
State administration area	1372	190	1	3.20	1.3
Institute/research center/higher education institution	1417	145	5	3.22	1.4
Foreign investment area	1423	139	5	3.22	1.4
International organizations, non-governmental organizations (NGOs)	1381	181	4	3.30	1.5

However, the thematic analysis of the participants' self-initiated responses reveals a more authentic landscape of the labor market with occupations clustered into eight categories. Diplomacy/international relations topped the list with 252 recommendations, followed by tourism (205), international economics/ business (196), and journalism/media (196). Groups with fewer opportunities included: translation/interpretation (164), education (143), international studies/research (106), management/human resources/administration (54), and non-governmental organizations (NGOs) and non-profit organizations (NPOs) (33). The results corroborate the closed-ended item regarding diplomacy/international relations, yet contradict non-governmental and non-profit organizations with fewer actual opportunities. Meanwhile, they also indicate more actual demands for tourism, international business/economics, and journalism/media sectors (Table 30).

Table 30

Occupational Fields Suitable for GIC Graduates as Recommended by Participants

Suitable career fields for GIC graduates	Recommendations
International Tourism/Hospitality	205

Diplomacy/ International Relations	252
International Economics/Business	196
PressMedia	196
Management/Human Resources/Administration	54
Education	143
Translation	164
International Studies/Research	106
International Organizations/Non-Governmental Organizations (NGOs)/ Non-Profit Organizations (NPOs)	33
Other	50
Total	**1399**

In general, with a globalized worldview, GIC graduates can expand their career opportunities in international environments such as diplomacy/international relations, international tourism/hospitality, international economics/business, and international journalism/media. Besides traditional professions like teaching and translation/interpretation, GIC graduates can assume many strategic positions within businesses, acting as a bridge for international business transactions, negotiating with foreign partners, localizing products, and operating international supply chains/platforms. To fulfil these responsibilities, besides knowledge and skills in languages, cultures, and translation/interpretation, students need to be equipped with a global cultural mindset, effective communication in multicultural environments, cultural empathy, conflict resolution, and negotiation skills. Below are some valuable suggestions regarding suitable career fields and the necessary competencies for GIC graduates.

*With a foundation in GIC, graduates are no longer limited to traditional translation and interpretation jobs but can **take on many strategic positions** in modern businesses. They can develop careers in multinational human resource management, where **cultural empathy skills** help build an inclusive and cohesive work environment. In Marketing and External Communications, sensitivity to regional psychology helps graduates become **global campaign coordinators**, ensuring brand messages do not violate sensitive cultural barriers. Furthermore, positions such as International Project Coordinator, Foreign Investment Consultant, or working for non-governmental organizations (NGOs) are also lucrative opportunities, as these jobs require **negotiation skills and conflict resolution** between parties with **differing values**. (Participant 53)*

*Graduates trained in globalization and intercultural communication can work in a wide range of fields thanks to their **language skills, cultural understanding, and global mindset**. Career opportunities span from international media, editing and content creation, PR, marketing, and international business to education, study abroad consulting, and socio-cultural research. Additionally, graduates can join non-governmental organizations, diplomatic agencies, and take on roles in project coordination or international relations. Fields such as tourism and hospitality, import/export, global human resources, **localized product development,** and **international platform operation** in the technology industry are also well-suited. Overall, careers for this group of graduates require **effective communication***

skills, an understanding of the international context, and the ability to work in a multicultural environment*. (Participant 961)*

What are recommended for the GIC curriculum development?

The thematic analysis of approximately 500 suggestions for GIC curriculum development from the final open-ended question yields three main themes: (a) core knowledge/ attributes, (b) specialized skills/competencies, and (c) curriculum implementation.

Core knowledge/attributes. The program should equip students with a rich worldview encompassing culture, international economics, international trade, politics, and global issues. The curriculum needs to be sensitive to and update the changing national, regional, and global landscapes, thereby corroborating and anticipating global trends. Moreover, in-depth specialized knowledge becomes more significant than ever as the international market demands competitive, high-quality human resources with standardized expertise and professionalism. Some suggested specializations include political science, global communication, international crisis management, negotiation in multinational environments, business, marketing, accounting, recruitment, human resources, and artificial intelligence applications. Moreover, learners need to pursue lifelong learning, constantly cultivating core attributes such as autonomy, self-management, dynamism, flexibility, quick adaptation, competitiveness, and global thinking.

Table 31.

Core Knowledge, Attributes to be Developed

General comments	Subthemes	Quantity
Knowledge	• General knowledge of worldview	7
	• Adapting to global development trends	5
	• Specialization requiring further training	8
Attributes	• Lifelong learning	7
	• Autonomy, self-management	2
	• Dynamism, flexibility, quick adaptation, competitiveness	4
	• Global thinking, global citizenship	4

Skills/competencies. While theory is undeniably important in developing abstract thinking at the university level, the vast majority of opinions emphasize PRACTICALITY: reality, practice, application, and competitiveness. This means that the GIC program must be closely linked to real-world experience, train practical skills that are highly applicable in real life, enhance enterprise practice, and improve competitive edge in the ever-changing international context. Four groups of skills/competences are summarized in

Table 32Error! Reference source not found.: intercultural language/communication competences, soft skills/critical thinking skills, digital competences, and professional skills.

The most fundamental core competency for GIC graduates is language proficiency and intercultural communication, with a total of 41 comments. Specifically, students need to be able to use English effectively and acquire internationally recognized certificates. Furthermore, multilingual proficiency (English, Russian, Chinese, Japanese, Korean, Arabic) provides a significant competitive advantage in a multinational environment. Language skills should not be limited to basic communication but should be applied in specialized fields such as logistics, finance, tourism, construction, and information technology. To nurture more demanding intercultural communication skills, students need to develop natural listening and speaking reflexes in international environments or in foreign joint ventures, where they may adopt new languages and (corporate) cultures.

Soft skills/critical thinking skills rank as the second most frequently mentioned skill group, with 23 comments. Soft skills can be interpreted as effective communication in the workplace, teamwork, adaptability and integration into a multinational community, the ability to continuously learn and develop oneself, and occupational soft skills. Furthermore, in a dynamic, high-pressure work environment, students need to cultivate critical thinking and problem-solving skills to improve productivity and work efficiency.

The third group is digital competence, with 18 contributions. In the era of digital transformation, when the world's businesses are operated by high-speed internet, high-tech platforms, and artificial intelligence, pioneering and adopting information technology in training and work is a vital requirement for development. Students desire more exposure to effective online communication platforms, virtual reality in experiential learning, and practice with specialized software applications in language, translation, and project management to enhance their digital competency to address workplace requirements.

The final group is professional skills. Although mentioned in only nine comments, this seems to be the most important skill set in career development because of its high applicability and relevance to businesses' needs, ensuring that graduates can quickly adapt to job positions after recruitment. Specialized skills in tourism, communication, marketing, journalism, business, import/export, technology, multinational teamwork, and intercultural communication should be trained more professionally in partnership with enterprises and foreign organizations.

Therefore, the GIC program needs to be designed with a focus on enhancing employability skills, encompassing four core skill/competence groups: standardized language proficiency with in-depth expertise and intercultural communication in a multinational environment; soft skills and thinking skills (critical thinking, problem-solving, adaptability); digital and artificial intelligence competences; and professional skills relevant to the training specialization.

Table 32

Employability Skills/Competences

Employability skills/Competences	Subthemes	Quantity
Skills	• Theory and practice	2
	• Reality/Practice	12
	• Applicability	9
	• Practical competition	2
Intercultural communication skills	• Standardized language proficiency	12
	• Multilingual proficiency	10
	• Highly specialized knowledge	4
	• Intercultural communication skills	15
Employability skills	• Soft skills/Critical thinking skills	23
	• Digital competences	18
	• Professional skills	9

Curriculum. The theme curriculum is arranged into three subthemes: orientation, scale, and quality of training (52 comments); program content and teaching-learning methods (28 comments); and training methods (63 comments).

Regarding training orientation, scale, and quality, GIC curriculum development is necessary to diversify training programs and address the human resource needs. The new program should shift from a language-focused mindset to training employability skills for a global environment, where students need to be equipped with in-depth knowledge of the culture, economy, politics, and society of other countries, specialized competences relevant to their profession, and the ability to network globally. Some experts and employers believe that today's recruitment trends require both foreign languages and professional skills, suggesting a dual-degree model with majors and minors, foreign languages and a second pathway, such as business, marketing, technology, law, or commerce, to enhance graduates' competitiveness in the labor market. Accordingly, participants agreed that the scale of GIC training needs to be expanded by increasing quotas, opening multidisciplinary classes, professional courses, and potentially integrating with existing training programs at other educational institutions from high school to university levels, including international schools. The intake and outcome quality are also crucial to generating a skilled and intellectual workforce for national development. This requires investment in human and material resources to develop the teaching staff and take care of students' well-being.

The program content and teaching methodology subtheme proposes innovating the curriculum to increase its applicability and practicality, developing soft skills and digital competencies, diversifying training content, and investing in digital technology to improve teaching quality. Furthermore, the curriculum needs to be refined and updated, with thematic talks on globalization, intercultural communication, business administration, and geopolitics; experiential learning at international enterprises; and participation in projects and activities with foreign partners. The training methodology should emphasize professionalism, systematic approaches, and transparent, effective quality assurance processes. Teaching methods should also be more engaging, creative, and diverse, incorporating simulated scenarios, real-world projects, business case studies, and internships at multinational companies and international organizations.

The training methodology highlights interdisciplinary/multidisciplinary, corporate, and intercultural communication skills training. Interdisciplinary and multidisciplinary training are becoming an urgent societal need as employers require language skills combined with professional expertise. Therefore, language training should be supported by employability skill development to generate a "multi-skilled" workforce with high adaptability to dynamic work environments. Corporate training is also highly recommended as it offers authentic work experience and helps students avoid “job shocks” upon commencement. The curriculum planners need to work in parallel with potential businesses with recruitment needs or a physical presence in Vietnam to fulfil their needs and cooperate with FDI companies to create more internships and job opportunities for graduates. This corporate training should be a mandatory assessment component and conducted rigorously to facilitate students' accumulation of work experience and their confident integration into international environments. At a higher level of cooperation, businesses and employers should be invited to participate in providing feedback and evaluating the curriculum. Eventually, the training of intercultural communication skills, an undeniable strength of VNU-ULIS in the era of integration and development, needs to be further developed by enhancing intercultural communication skills, moving beyond theoretical understanding to application and creativity through realistic simulations to strengthen international academic exchange and connect students with international friends, tourists, and expats working in Vietnam. "Cultural empathy" is perhaps the most important attribute, opening up opportunities for collaboration and development on many levels.

Table 33

GIC Training Program

GIC training program	Subthemes	Quantity
Orientation, scale, and quality of training	• Training orientation	21
	• Training scale	15
	• Training quality	16
Program content and teaching methods	• Program content	12
	• Teaching methods	16
Training methods	• Interdisciplinary/Multidisciplinary training	23

	• Corporate training/labor market/ corporate internships/training partnerships with businesses	23
	• Enhancing international communication skills	17

Below is a summary of typical experts/researchers and employers' contributions regarding the vision and GIC development strategies. The most prominent training strategy is a shift in focus from "language proficiency" to "context mastery," with languages becoming the foundation for developing cultural intelligence, global thinking skills, management, digital competences, and communication skills. GIC graduates will be able to analyze global capital flows, effectively connect links in international supply chains, and engage in lifelong learning to adapt to the fluctuating labor market. Intercultural communication competences are increasingly important, as they enable students to understand "cultural symbols," thereby effectively explaining and bridging differences. Furthermore, communication needs to go hand in hand with employability skills and adaptability. Thus, students need a multifaceted, insightful perspective of the target occupations and should not be afraid to engage with real-world situations to improve themselves. Accordingly, interdisciplinary/multidisciplinary curricula should be developed to provide practical skills, real-world experience, soft skills, global competencies, and digital technologies.

In the context of artificial intelligence (AI) gradually replacing basic language tasks, the focus of foreign language training needs to shift from ***"verbal fluency"*** *to* ***"context mastery."*** *Current curricula should be designed as* ***an integrated model****, where foreign languages play a foundational role, combined with essential supporting skills such as* ***managerial thinking, digital skills****, and especially* ***cultural intelligence (CQ)****. Society needs not only fluent foreign language speakers but also personnel capable of* ***analyzing the movement of global capital flows and effectively connecting links in international supply chains****. Therefore, universities need to enhance practical application through internships* ***at FDI enterprises****, while equipping students with a lifelong learning mindset to* ***adapt to the rapidly changing labor market****. (Participant 51)*

In the context of deep integration, the demand for human resources extends beyond mere foreign language proficiency; it also requires individuals ***with intercultural communication skills, global thinking, and an understanding of international social and political contexts****. Foreign language undergraduate programs should broaden their scope to include skills such as* ***multicultural context analysis, international negotiation, multilingual organizational management****, and* ***digital skills in multimedia communication****. This enables students to not only translate languages but also interpret* ***"cultural symbols,"*** *meaning* ***they can effectively understand, interpret, and connect with differences.*** *Training also needs to be more* ***flexible*** *with practical modules, connections to businesses and international organizations, and exchange programs. At the same time, the program should integrate interdisciplinary knowledge, such as international business, project management, communication, and information technology, to meet the diverse career requirements of the modern job market. (Participant 539)*

In summary, the proposals for developing the GIC program affirm its urgency at VNU-ULIS, creating possibly more lucrative job opportunities and aligning with the nation's more intensive international integration. The program illuminates a shift in training orientation from "language fluency" to "context mastery" with a competence-based curriculum covering four core skill/competency groups: intercultural language/communication competence, soft skills, digital/ artificial intelligence competence, and professional skills. Thus, more investment is required to improve quality and expand scale. While the GIC curriculum can be implemented at the school level, this orientation should be integrated into existing training programs inside and outside VNU-ULIS. The training content needs to be comprehensive, up-to-date, and flexible to meet the needs of society and employers. Educational methods should be engaging, creative, and highly applicable through simulation exercises, case studies, collaborative projects with businesses, and international exchanges. Training modes should aim towards interdisciplinary/multidisciplinary, corporate-based training, and intercultural communication skills development.

4. Discussions and conclusion

The feasibility study on launching a new GIC program at VNU-ULIS was conducted with the research team's strong commitment. The documentary analysis of the state and government's policies and the existing curricula, together with the large-scale needs assessment with both quantitative and qualitative data, has encapsulated the dynamic interplay among the quintuple helices of the multi-layered ecosystem. After approximately six months of data collection through the Cognito survey questionnaire, the research obtained 1562 valid responses and more than 500 suggestions for GIC curriculum development. Based on the conceptual framework of a multi-tiered ecosystem (Bronfenbrenner, 1979) and the tri-dimensional needs analysis of PSA, TSA, and CNA, the research results were arranged and presented scientifically in the four levels of the ecosystem.

5. Summary of research results

RQ1. At the macro layer, the State and Government's policies and strategies promote globalization and comprehensive international integration in education to enhance Vietnam's role and its position in the international arena.

RQ2. At the meso layer, VNU-ULIS has a high capacity to meet the requirements for launching the GIC program in terms of curriculum (92/162 highly compatible subjects), human resources (45 lecturers and other functional departments), and facilities (over 170 multimedia classrooms, digital platforms, and a dynamic, modern ULIS ecosystem).

RQ3. At the exo and micro levels, a large-scale needs survey ($N = 1562$) with 5 relevant target groups - high school students, university students, workers, employers, and experts/researchers - showed that 57% were interested/very interested in the training program and career opportunities could expand in tourism/hospitality, diplomacy/international relations, economics/international business, and journalism/communication, instead of being limited to the traditional fields of education and translation/interpretation.

RQ4. The thematic analysis of multiple stakeholders' recommendations has yielded a competence-based framework for GIC curriculum development consisting of linguistic and intercultural competence, digital competence, soft skills, and professional skill

Implications for developing the GIC program

GIC orientation. As the research results indicate, the GIC orientation is consistently aligned among all four levels in the human resource development ecosystem. International

integration is perceived as the driving force for the next stage of national development and modernization. This orientation has emerged in VNU-ULIS's existing curricula, with highly compatible subjects accounting for 56.8% of the examined. Research found that GIC compatibility is most discernible in language and culture knowledge (94.6%), general knowledge (60.1%), and less so in the specialized knowledge block (38.1%). Therefore, the university needs to refine the existing compatible subjects to align the teaching content with a GIC orientation and develop more GIC-specialized subjects, such as international economics, trade negotiation, and international event analysis, to increase the professional knowledge content of the university's overall orientation.

The GIC orientation cannot only be integrated into existing VNU-ULIS curricula, but also transferred to other VNU members, or developed into an independent curriculum in an attempt to internationalize universities (Rizvi, 2007), as internationalization of higher education has recently been regarded as a game changer not only in developed but also in emerging and developing societies (de Wit, 2020). In other words, the GIC orientation resembles "integrating an international/intercultural dimension into the teaching, research, and service functions of the institution" (Knight, 1994, as cited in OECD, 1999, p.16) to shape a global perspective (van Kampen et al., 2021). If intercultural communication, specialized English, and negotiation skills courses are embedded in high-quality curricula of VNU member universities, such as the University of Economics and Business, University of Engineering and Technology, University of Law, University of Education, University of Medicine and Pharmacy, the Interdisciplinary Training Faculty, and the International School, it can generate a more competitive workforce for national development. Therefore, global connectivity should be incorporated into the curricula of other universities, academies, and even high schools. In parallel with the policy of making English the second language in Vietnam, educational institutions need to adopt policies for recruiting global citizens who can proficiently use at least one foreign language for international communication in tourism, business, trade, educational and cultural exchanges in their home country, abroad, or in the internet environment.

Developing a GIC curriculum. The development of a GIC program needs to be grounded in intensive research about the market economy and its operation under the influence of scientific, technological, socioeconomic, and geopolitical factors. The research results sketch a competence-based framework for curriculum development, encompassing language and intercultural communication competence, soft/critical thinking skills, digital/artificial intelligence competence, and professional skills. The objectives and program learning outcomes need to adhere to this framework and reflect the national demand for economic and sociocultural development. Furthermore, the curriculum needs to ensure practicality and applicability, not only equipping students with appropriate knowledge, skills, and attitudes, but also developing the competencies and attributes of future global citizens who can quickly adapt to a changing world. This orientation is aligned with "contemporary appreciation of the fluid nature of cultural and workplace identities" (Dai, 2024, p. 452) and inspires the implementation of the intercultural dimension in language teaching (Huang & Fang, 2023).

A new GIC curriculum will converge all the best fits into a single training program and reinforce it with newly designed courses to facilitate international integration and cooperation. The courses in the GIC program can be inherited from 92 highly compatible existing subjects, with adaptations and updates in objectives and content. The GIC program should also provide comprehensive interdisciplinary/multidisciplinary training in law, business, intercultural communication, political event analysis, and essential 21st-century competencies, such as digital skills, critical thinking skills, adaptability, and professional skills. Additionally, the

university needs to research and develop more practical, applicable specialized courses, such as *Negotiation Art, International Economics, International Law, and International Event Analysis,* to enhance learners' competence in international integration.

Specialized and corporate training in collaboration with enterprises and FDI corporations will minimize workplace retraining and improve GIC graduates' competitiveness with their counterparts from other leading institutions, such as the University of Foreign Trade, the Diplomatic Academy, National Economics University, and Vietnam-Japan University, which offer similar majors. The training methods for the GIC program also need to be innovative and engaging. This could include enhancing practical experience, teaching through simulated scenarios, inviting experts and recruiters to share their real-world experiences, and placing greater emphasis on developing language skills, multilingualism, and intercultural communication. Furthermore, the university could consider expanding its collaboration with member universities of Vietnam National University, Hanoi, international counterparts, and FDI enterprises to offer high-quality, internationally standardized programs to accredit the international work skills of GIC graduates.

Limitations and research recommendations

The research was conducted urgently and received enthusiastic support from many stakeholders. While snowball sampling yielded a large amount of both quantitative and qualitative data, it is a non-probability sampling method, requiring caution in generalizing the research results. Besides the voluntary participation of a large number of high school students, undergraduates, and working professionals, the research team encountered difficulties in approaching experts/researchers and employers in this relatively new field, relying merely on personal connections. Further research, with support from the VNU-ULIS network in sending formal invitations to interviewees or organizing conferences and seminars, could gather more constructive feedback from experts, researchers, and employers. Studies on the internationalization of current training programs at VNU-ULIS and the integration of intercultural communication skills into elite training programs of VNU member universities could be potential research topics for experts interested in this field.

Conflict of interest

The authors declare that there was no conflict of interest.

Sponsorship

This research was funded by VNU University of Languages and International Studies (VNU-ULIS) in project No. N.25.21.

AI declaration

Copilot and Gemini were adopted to brainstorm ideas and synthesize the State and Government's policies. The research team conducted the data collection, analysis, and writing up. The final scientific research report was translated into English by Google Translate. The transformation and refinement of the report into an article were completed by the research team, with language polish supported by Grammarly.

References

Agrawal, R., Golshan, B., & Papalexakis, E. (2016). Toward data-driven design of educational courses: A feasibility study. *Journal of Educational Data Mining*, *8*(1), 1-21

Andi, K., & Arafah, B. (2017). Using needs analysis to develop English teaching materials in initial speaking skills for Indonesian college students of English. *The Turkish Online Journal of Design, Art and Communication* (10), 419-437

Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Medical Research Methodology*, *10*(1), 67. https://doi.org/10.1186/1471-2288-10-67

Arnó-Macià, E., Aguilar-Pérez, M., & Tatzl, D. (2020). Engineering students' perceptions of the role of ESP courses in internationalized universities. *English for Specific Purposes*, *58*, 58–74. https://doi.org/10.1016/j.esp.2019.12.001

Ban Chấp hành Trung ương. (2025, October 2). *Nghị quyết 71-NQ/TW 2025 đột phá phát triển giáo dục và đào tạo*. https://thuvienphapluat.vn/van-ban/Giao-duc/Nghi-quyet-71-NQ-TW-2025-dot-pha-phat-trien-giao-duc-va-dao-tao-670683.aspx

Báo điện tử Chính phủ. (2026, June 1). *Toàn văn: Nghị quyết 57-NQ/TW về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia*. https://xaydungchinhsach.chinhphu.vn/toan-van-nghi-quyet-ve-dot-pha-phat-trien-khoa-hoc-cong-nghe-doi-moi-sang-tao-va-chuyen-doi-so-quoc-gia-119241224180048642.htm

Báo Nhân Dân điện tử. (2025, June 2). Thực thi mô hình hợp tác "3 nhà." *Báo Nhân Dân điện tử*. https://nhandan.vn/thuc-thi-mo-hinh-hop-tac-3-nha-post884137.html

Benesch, S. (2001). *Critical English for Academic Purposes: Theory, politics, and practice*. Routledge. https://doi.org/10.4324/9781410601803

Bộ Chính trị. (2021, March 22). *Chiến lược phát triển kinh tế - xã hội 10 năm 2021-2030*. https://tulieuvankien.dangcongsan.vn/ban-chap-hanh-trung-uong-dang/dai-hoi-dang/lan-thu-xiii/chien-luoc-phat-trien-kinh-te-xa-hoi-10-nam-2021-2030-3735

Bộ Chính trị. (2026, January 16). *Nghị quyết số 80-NQ/TW ngày 07/01/2026 của Bộ Chính trị về phát triển văn hóa Việt Nam*. tulieuvankien.dangcongsan.vn. https://tulieuvankien.dangcongsan.vn/he-thong-van-ban/van-ban-cua-dang/nghi-quyet-so-80-nq-tw-ngay-07-01-2026-cua-bo-chinh-tri-ve-phat-trien-van-hoa-viet-nam2.html

Bộ Giáo dục và Đào tạo. (2025, January 2). *Chiến lược phát triển giáo dục đến năm 2030, tầm nhìn đến năm 2045*. baochinhphu.vn. https://baochinhphu.vn/chien-luoc-phat-trien-giao-duc-den-nam-2030-tam-nhin-den-nam-2045-102250102165657226.htm

Boroujeni, S. A., & Fard, F. M. (2013). A needs analysis of English for Specific Purposes (ESP) course for adoption of communicative language teaching: A case of Iranian first-year students of educational administration. *International Journal of Humanities and Social Science Invention*, *2*(6), 35–44.

Borrero, J. D., & Yousafzai, S. (2024). Circular entrepreneurial ecosystems: A quintuple helix model approach. *Management Decision*, *62*(13), 141–177. https://doi.org/10.1108/MD-08-2023-1361

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., Fabrizio, C., & Fernandez, M. (2009). How we design

feasibility studies. *American Journal of Preventive Medicine*, *36*(5), 452–457. https://doi.org/10.1016/j.amepre.2009.02.002

Carayannis, E., & Campbell, D. F. J. (2012). Triple helix, quadruple helix and quintuple helix and how do knowledge, innovation and the environment relate to each other?: A proposed framework for a trans-disciplinary analysis of sustainable development and social ecology. In *Sustainable Policy Applications for Social Ecology and Development* (pp. 29–59). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-4666-1586-1.ch004

Chambers, F. (1980). A re-evaluation of needs analysis in ESP. *The ESP Journal*, *1*(1), 25–33. https://doi.org/10.1016/0272-2380(80)90007-4

Cloitre, A., Dos Santos Paulino, V., & Theodoraki, C. (2023). The quadruple/quintuple helix model in entrepreneurial ecosystems: An institutional perspective on the space case study. *R&D Management*, *53*(4), 675–694. https://doi.org/10.1111/radm.12547

Dai, D. W. (2024). Interactional Competence for professional communication in intercultural contexts: Epistemology, analytic framework and pedagogy. *Language, Culture and Curriculum*, *37*(4), 435–455. https://doi.org/10.1080/07908318.2024.2349781

Dai, D. W. (2026). What do second language speakers really need for real-world interaction? A needs analysis of L2 Chinese interactional competence. *Language Teaching Research*, *30*(2), 655–692. https://doi.org/10.1177/13621688221144836

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2025). *Đề án chương trình đào tạo trình độ đại học: Toàn cầu hóa và giao tiếp xuyên văn hóa (tiếng Anh)*. Mã số: 7310206

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2017). *Chương trình đào tạo đại học ngành Ngôn ngữ Anh*. https://daotao.ulis.vnu.edu.vn/chuong-trinh-giao-duc-dai-hoc-nganh-ngon-ngu-anh/

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2019). *Chương trình đào tạo chuẩn trình độ đại học ngành: Sư phạm tiếng Anh Mã số: 7140231*.Signed.Signed.1.-Sư-phạm-Tiếng-Anh.pdf

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2022). *Đề án chương trình đào tạo thí điểm văn hóa và truyền thông xuyên quốc gia trình độ đại học của Trường Đại học Ngoại ngữ*

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2023). *Chương trình đào tạo chuẩn trình độ đại học ngành VHTTXQG (Mã số: 7220212QTD)*.Signed.Signed.15.-Văn-hóa-và-Truyền-thông-X.pdf

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2025a). *Chương trình đào tạo Giảng dạy tiếng Việt như một ngoại ngữ (ngành Tiếng Việt và Văn hóa Việt Nam)*. https://daotao.ulis.vnu.edu.vn/chuong-trinh-dao-tao-nganh-tieng-viet-va-van-hoa-viet-nam-ap-dung-tu-khoa-qh2025/

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2025b). *Công tác mở ngành, chuyên ngành đào tạo tại Trường Đại học Ngoại ngữ - ĐHQGHN giai đoạn 2025-2030*.

Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. (2025c, August 29). *Công bố các báo cáo tự đánh giá chương trình đào tạo năm 2025*. https://ulis.vnu.edu.vn/cong-bo-cac-bao-cao-tu-danh-gia-chuong-trinh-dao-tao-nam-2025/

Đảng Cộng sản Việt Nam. (2021, February 26). *Nghị quyết Đại hội đại biểu toàn quốc lần thứ XIII của Đảng*. tulieuvankien.dangcongsan.vn.

https://tulieuvankien.dangcongsan.vn/ban-chap-hanh-trung-uong-dang/dai-hoi-dang/lan-thu-xiii/nghi-quyet-dai-hoi-dai-bieu-toan-quoc-lan-thu-xiii-cua-dang-3663

de Wit, H. (2020). Internationalization of higher education: The need for a more ethical and qualitative approach. *Journal of International Students*, *10* (1), i–iv. https://doi.org/10.32674/jis.v10i1.1893

Dewi, H. H., Hidayatulloh, S. M. M., Sukarno, S., Lestari, A. E., Dewi, I. L., & Ciptaningrum, D. S. (2023). English materials development for an undergraduate communication study program: A need analysis in Indonesian context. *LLT Journal: A Journal on Language and Language Teaching*, *26*(1), 69–91. https://doi.org/10.24071/llt.v26i1.5208

Dou, A. Q., Chan, S. H., & Win, M. T. (2023). Changing visions in ESP development and teaching: Past, present, and future vistas. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1140659

Hà, H. T. N., & Cúc, T. T. K. (2025). Hợp tác ba nhà theo mô hình triple helix tại trường Đại học Tân Trào: Thực trạng và giải pháp. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*, 204–204. https://doi.org/10.59266/houjs.2025.944

Huang, W., & Fang, F. (2023). EMI Teachers' perceptions and practices regarding culture teaching in Chinese higher education. *Language, Culture and Curriculum*, *36*(2), 205–221. https://doi.org/10.1080/07908318.2022.2115056

Hutchinson, T., & Waters, A. (1987). *English for Specific Purposes*. Cambridge University Press. https://doi.org/10.1017/CBO9780511733031

Kaewpet, C. (2009). A framework for investigating learner needs: Needs analysis extended to curriculum development. *Electronic Journal of Foreign Language Teaching*, *6*(2), 209–220.

Lan P. T. (2025). Đẩy mạnh phát triển khoa học - công nghệ, đổi mới, sáng tạo. *Quản lý nhà nước*, (351), 14–19. https://doi.org/10.59394/qlnn.351.2025.1138

Lê Đ. B., Phan H., Nguyễn A. T., & Ngô T. C. V. (2025). Mô hình hợp tác nhà trường – doanh nghiệp trong đào tạo nhân lực bán dẫn: Kinh nghiệm quốc tế và hàm ý cho Việt Nam. *Tạp chí Nghiên cứu Khoa học và Phát triển*, 99–108. https://doi.org/10.58902/tcnckhpt.v4i4.340

Li, J. (2014). Needs analysis: An effective way in business English curriculum design. *Theory and Practice in Language Studies*, *4*(9), 1869–1874. https://doi.org/10.4304/tpls.4.9.1869-1874

Long, M. H. (Ed.). (2005). *Second language needs analysis*. Cambridge University Press.

https://doi.org/10.1017/CBO9780511667299

Morawska-Jancelewicz, J. (2022). The role of universities in social innovation within the quadruple/quintuple helix model: Practical implications from Polish experience. *Journal of the Knowledge Economy*, *13*(3), 2230–2271. https://doi.org/10.1007/s13132-021-00804-y

Chính phủ Việt Nam. (2018). Nghị định 86/2018/NĐ-CP quy định về hợp tác, đầu tư của nước ngoài trong lĩnh vực giáo dục, Chính phủ Việt Nam. https://datafiles.chinhphu.vn/cpp/files/vbpq/2018/06/86.signed.pdf

OECD. (1999). *Quality and Internationalisation in Higher Education*. OECD. https://doi.org/10.1787/9789264173361-en

Orsmond, G. I., & Cohn, E. S. (2015). The distinctive features of a feasibility study: Objectives and guiding questions. *OTJR: Occupation, Participation and Health*, *35*(3), 169-177. DOI: 10.1177/1539449215578649

Rahman, M. (2015). English for Specific Purposes (ESP): A holistic review. *Universal Journal of Educational Research*, *3*(1), 24–31. https://doi.org/10.13189/ujer.2015.030104

Richards, J. C. (2009). *Curriculum development in language teaching*. Cambridge University Press.

Rizvi, F. (2007). Internationalization of curriculum: A critical perspective. In M. Hayden, J. Levy, & J. Thompson (Eds.), *The Sage Handbook of Research in International Education* (pp.390–403). Sage Publications Ltd. https://doi.org/10.4135/9781848607866.n33

Thạch T. Đ. N. (2022). Quan hệ hợp tác giữa nhà trường và doanh nghiệp trong mô hình đào tạo định hướng ứng dụng. *Tạp chí Khoa học Công nghệ - Đại học Hòa Bình.* 115–121. https://daihochoabinh.edu.vn/quan-he-hop-tac-giua-nha-truong-va-doanh-nghiep-trong-mo-hinh-dao-tao-dinh-huong-ung-dung

Thi P. Đ., & Bình L. T. T. (2026). Xây dựng cơ chế ba nhà trong hợp tác, đối tác nghiên cứu khoa học ở Trường Đại học Sân khấu - Điện ảnh Hà Nội. *Tạp chí Nghiên cứu Sân khấu - Điện ảnh*, 49. https://doi.org/10.70494/2354-0680.282

Thủ tướng Chính phủ. (2025). *Quyết định số 2371/QĐ-TTg của Thủ tướng Chính phủ: Phê duyệt Đề án "Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học giai đoạn 2025 - 2035, tầm nhìn đến năm 2045."* http://vanban.chinhphu.vn/?pageid=27160&docid=215728

Thủ tướng Chính phủ. (2021). *Quyết định số 1909/QĐ-TTg của Thủ tướng Chính phủ: Về việc phê duyệt Chiến lược phát triển văn hóa đến năm 2030.* http://chinhphu.vn/?pageid=27160&docid=204463&tagid=6&type=1

Tight, M. (2022). Internationalisation of higher education beyond the West: Challenges and opportunities – the research evidence. *Educational Research and Evaluation*, *27*(3–4), 239–259. https://doi.org/10.1080/13803611.2022.2041853

van Kampen, E., Meirink, J., Admiraal, W., & Berry, A. (2021). Characterising integrated content-language pedagogies of global perspectives teachers in Dutch bilingual schools. *Language, Culture and Curriculum*, *34*(1), 18–34. https://doi.org/10.1080/07908318.2020.1732999

Vân V. L., Hà N. T. T., & Anh N. T. K. (2026). Ứng dụng mô hình hợp tác "ba nhà" trong giảng dạy trí tuệ nhân tạo: Nghiên cứu thí điểm tại các trường trung học phổ thông ở tỉnh Lâm Đồng. *Tạp chí Giáo dục*, *26*(8), 90–96. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/5493/

Viện Chiến lược Phát triển. (2024, May 21). *Báo cáo Việt Nam 2045: Xu hướng kinh tế toàn cầu và hàm ý chính sách cho Việt Nam*. Viện Chiến lược và Chính sách Kinh tế - Tài chính. https://mpi.gov.vn/tin-tuc/bao-cao-viet-nam-2045-xu-huong-kinh-te-toan-cau-va-ham-y-chinh-sach-cho-viet-nam-3280.html

#151 THE USE OF ARTIFICIAL INTELLIGENCE IN ACADEMIC WRITING AND ITS RELATIONSHIP WITH SELF-REGULATED LEARNING: A SYSTEMATIC REVIEW

Thach Huong Giang[1], Cao Thanh Cong[2], Trinh Hong Linh[3]

Abstract

The increasing use of artificial intelligence (AI), particularly generative AI tools, has reshaped how students approach academic writing in English as a foreign language (EFL) contexts. While AI can enhance writing performance by providing immediate and personalized feedback, its relationship with self-regulated learning (SRL) remains unclear. This systematic review synthesizes recent empirical studies to examine how AI use is associated with learners' SRL in academic writing. The findings suggest that AI can support writing quality, motivation, and efficiency when used strategically. However, excessive or uncritical use may limit learners' engagement in higher-order cognitive processes, including critical thinking and self-monitoring. The review also highlights the importance of learners' interaction with AI-generated feedback in shaping learning outcomes, proposes concrete pedagogical strategies for classroom integration, and identifies a lack of empirical evidence in specific contexts, particularly in Vietnam. The study underscores the need for pedagogically guided AI integration to promote both effective writing development and sustainable self-regulated learning.

***Keywords*:** artificial intelligence, AI, academic writing, self-regulated learning, SRL, EFL

1. Introduction

Academic writing in English is a complex, essential, and highly integrated skill for university students because it serves not only as a means of demonstrating knowledge but also as a medium through which learners develop critical thinking, reasoning, and the ability to express ideas logically (Gao et al., 2025). Recent advances in artificial intelligence (AI), particularly Generative AI, have significantly transformed the ways students approach and practice academic writing in English (Huang & Mizumoto, 2024). GenAI tools can provide immediate feedback on grammar, vocabulary, coherence, and text structure, while also assisting learners in optimizing the writing process, improving content quality, and enhancing overall learning efficiency (Chan & Lee, 2025; Kim et al., 2025). Alongside these technological changes, self-regulated learning (SRL) - the ability of learners to plan, monitor, and evaluate their own learning (Zimmerman, 1986) - continues to play a key role in successful academic writing.

Despite these benefits, the growing use of AI in academic writing has raised concerns among educators and researchers. One of the most frequently discussed issues is students' potential overreliance on AI tools (Marzuki et al., 2023). Existing studies indicate that although AI can improve the coherence and quality of writing work, it may simultaneously reduce students' independent thinking and creativity (Marzuki et al., 2023; Shi et al., 2025). Some learners have reported that frequent use of AI makes them less willing to write independently and more inclined to depend on AI-generated responses. Such concerns suggest that excessive

[1] *VNU University of Education; Email: giangthach0610@gmail.com*
[2] *Foreign Trade University; Email: caothanhcong2001@gmail.com*
[3] *University of Languages and International Studies, Vietnam National University; Email: thlinh@vnu.edu.vn*

reliance on AI may weaken learners' self-directed learning capacity, which is considered a fundamental element of sustainable academic development (Marzuki et al., 2023; Shi et al., 2025). As a result, understanding how AI influences students' self-regulated learning has become an important issue.

Despite growing interest in AI-assisted writing, the relationship between AI and self-regulated learning is still not well understood. Although the use of AI in academic writing has attracted increasing attention, existing studies have focused on describing the benefits of AI tools or learners' attitudes toward their use rather than examining their influence on self-regulated learning (Kim et al., 2025; Ozdere, 2025). Studies that discussed SRL often examined only one aspect, such as motivation, critical thinking, or self-assessment (Dazzeo, 2024; Gao et al., 2025; Huang & Mizumoto, 2024), rather than considering SRL as a complete learning process. Consequently, the current evidence remains fragmented, making it difficult to understand the overall impact of AI on students' self-regulation.

Another area that has received limited attention is how students interact with AI-generated feedback. Feedback is one of the main ways through which AI supports learning, yet relatively little is known about how students use the suggestions provided by AI. Some students critically evaluate AI feedback before revising their writing, whereas others accept AI suggestions with little reflection (Chan & Lee, 2025; Shibani et al., 2024). These different patterns of interaction may influence whether AI strengthens or weakens SRL. However, existing studies have not been systematically reviewed to explain these differences.

In addition, most research on AI-assisted academic writing has been conducted in Western countries and a small number of Asian contexts (Hornberger et al., 2025; Wang et al., 2024). Evidence from Vietnam remains very limited, even though generative AI is becoming increasingly popular in Vietnamese higher education. Because educational practices, English proficiency, and technology use vary across countries, findings from other contexts may not fully reflect the experiences of Vietnamese students. A systematic review is therefore needed to bring together existing evidence, identify common findings, and highlight areas where further research is required.

This systematic review aims to synthesize empirical studies on the use of AI in English academic writing, with a particular focus on its relationship with students' SRL and their interaction with AI-generated feedback. Specifically, the review has three objectives: (1) to examine how AI influences different aspects of self-regulated learning in academic writing, (2) to explore how students interact with AI-generated feedback, and (3) to identify gaps in the current literature, especially those related to the Vietnamese EFL context. The review addresses the following research questions:

RQ1. How does the use of AI influence students' self-regulated learning in academic writing?

RQ2. How do students interact with AI-generated feedback?

This review includes empirical studies published between 2023 and 2025, the period during which AI became widely adopted in education. Only English-language studies indexed in Scopus, Web of Science, ERIC, and Google Scholar were included. Qualitative, quantitative, and mixed-methods studies were eligible, whereas conceptual papers, technical reports, and studies unrelated to academic writing or learning were excluded. The review adopts Zimmerman's (1986) broad concept of SRL and synthesizes findings across different academic disciplines.

2. Methodology

2.1 Research design

This study employed a thematic literature review to synthesize empirical evidence on the influence of AI on students' SRL in academic writing and their interaction with AI-generated feedback. A thematic approach was considered appropriate because the review aimed to identify common themes, patterns, and gaps across existing studies rather than combine statistical results. Given the methodological diversity of the included studies, which comprised qualitative, quantitative, and mixed-methods research, the review focused on integrating findings conceptually to provide a comprehensive understanding of the topic and address the research questions.

2.2 Search strategy

Relevant studies were identified through searches in major academic databases, including Scopus, Web of Science, ERIC, and Google Scholar. The search process focused on studies published between 2023 and 2025, reflecting the period during which AI tools became widely adopted in educational settings.

A combination of keywords and Boolean operators was used, including “AI”, “generative AI”, “GenAI” OR “AI chatbot”, “self-regulated learning”, “academic writing” OR “EFL writing”, “AI feedback” OR “automated feedback”, and “feedback uptake” OR “student interaction”. These keywords were combined using AND and OR operators to retrieve studies relevant to the two research questions.

2.3 Inclusion and exclusion criteria

Studies were selected according to predefined inclusion and exclusion criteria to ensure the relevance and quality of the review. To be included, studies had to be published between 2023 and 2025, written in English, and focus on GenAI or AI-assisted writing tools within educational contexts. Eligible studies were required to examine topics related to self-regulated learning, academic writing, AI-generated feedback, or students’ interactions with AI feedback. In addition, only empirical studies employing qualitative, quantitative, or mixed-methods approaches were considered. To ensure academic rigor, the review included studies published in peer-reviewed journals, conference proceedings, or scholarly book chapters.

Studies were excluded if they focused solely on the technical development of AI systems without addressing educational applications, or if they addressed contexts unrelated to English-language learning or writing. Conceptual or theoretical papers that did not present empirical findings were also excluded. Duplicate records and publications with inaccessible full texts were removed during the screening process.

2.4 Study selection

The database search yielded 57 potentially relevant records. Among these, 47 investigated the relationship between AI and SRL, while 10 explored the concept of AI reliance or dependency. All records were assessed using the inclusion and exclusion criteria. As the screening and full-text eligibility assessment were conducted in a single stage, these processes are reported together, which is acknowledged as a limitation of the review methodology. Following the eligibility assessment, 37 records were excluded because they did not examine SRL or AI-generated feedback as a primary focus, or were not situated in an English-language

learning or academic writing context. Consequently, 20 empirical studies satisfied all inclusion criteria and were included in the thematic synthesis.

2.5 Data analysis

A thematic synthesis approach was adopted to analyze the selected studies. Information extracted from each article included publication details, research context, participants, methodology, AI tools used, and major findings.

The extracted findings were coded and grouped according to the research questions. For RQ1, themes focused on the influence of generative AI on dimensions of self-regulated learning, including motivation, self-monitoring, self-assessment, reflection, and autonomy. For RQ2, themes focused on learners' interaction with AI-generated feedback, including trust, evaluative judgment, feedback uptake, feedback rejection, and reflective engagement.

The synthesis process aimed to identify common patterns, areas of agreement, and emerging concerns across the reviewed literature, thereby providing a comprehensive understanding of the educational implications of generative AI in academic writing.

Table 1

Summary of studies included in the thematic synthesis.

Study	AI Tool / Focus	Design	Participants	Key Finding
Apriani et al. (2024)	AI chatbot	Mixed-methods	EFL students	writing performance, confidence, self-reported SRL
Wang et al. (2024)	ChatGPT	Mixed-methods	Postsecondary learners	92% positive perception; self-management/monitoring imbalance
Chang et al. (2023)	AI chatbot (design principles)	Conceptual/design-oriented	Not specified	AI supports SRL via goal setting, feedback, personalization
Wei (2023)	AI-supported platform	Mixed-methods	Not specified	motivation & SRL vs. traditional instruction
Dazzeo (2024)	AI self-assessment tool	Not specified	Students (writing course)	reflective thinking, purposeful revision
Wong & Viberg (2024)	ChatGPT	Two empirical studies	Not specified	Supports SRL in writing/reading; feedback lacks depth
Hapsari et al. (2023)	AI-mediated self-assessment	Not specified	University students	Trust & proficiency shape engagement
Yang et al. (2024)	Pigai (AWE system)	Exploratory	EFL writers	Feedback favors grammar; evolves toward self-regulated revision
Zhan & Yan (2025)	ChatGPT feedback	Not specified	Students	Domain-specific trust; superficial copying behavior
Thompson (2025)	AI personalized feedback	Not specified	Undergraduates	self-efficacy & performance; risk of reduced autonomy
Marzuki et al. (2023)	AI writing tools	Qualitative (teacher perspective)	EFL teachers	AI aids coherence but may reduce independent thinking

Shi et al. (2025)	ChatGPT vs. AWE	Comparative	L2 learners	Differing effects on ideal L2 writing self
Gao et al. (2025)	AI chatbot-integrated tasks	Exploratory	Not specified	Effects on self-efficacy & critical thinking
Kim et al. (2025)	GenAI writing tools	Not specified	Students	Perspectives on GenAI-assisted writing
Huang & Mizumoto (2024)	GenAI	Not specified	Students	motivation & writing self-efficacy
Chan & Lee (2025)	GenAI	Case study	Secondary school students	Balancing AI use and authenticity
Ozdere (2025)	ChatGPT, You.com	Not specified	EFL students	Effectiveness & grading consistency assessed
Shibani et al. (2024)	AI feedback (general)	Not specified	Students	Critical interaction with AI in assessment
Hornberger et al. (2025)	AI literacy (general)	Multinational survey	University students (DE, UK, US)	AI literacy patterns across contexts

Note. Fields marked "Not specified" reflect information not reported in the original studies as cited in this review.

3. Results

3.1. Influence of AI on students' self-regulated learning in academic writing (RQ1)

Evidence from the included studies suggests that AI can support students' self-regulated learning (SRL), although its benefits are not automatic (Apriani et al., 2024; Wang et al., 2024). Rather than serving only as a writing tool, AI is presented as a learning partner that can support several aspects of SRL, including goal setting, self-monitoring, self-evaluation, motivation, and reflective learning (Chang et al., 2023; Wei, 2023; Dazzeo, 2024; Wong & Viberg, 2024; Wang et al., 2024).

Chang et al. (2023) argue that AI chatbots are most beneficial when they encourage learners to set goals, evaluate their own work, and receive personalized support, rather than simply generating text or correcting errors. Their findings suggest that AI promotes self-regulation when it encourages students to reflect on their learning and adjust their writing strategies. In contrast, using AI only to produce answers or make corrections may encourage a more passive approach to learning.

The positive influence of AI on motivation and confidence is another consistent finding across the reviewed studies. Wei (2023) found that students using an AI-supported platform reported higher levels of motivation and self-regulation than those receiving traditional instruction. This improvement was associated with a learning environment that reduced writing anxiety and encouraged a growth mindset. Similarly, Apriani et al. (2024) reported improvements in students' writing performance, confidence, self-reported SRL, and ability to monitor their own progress after using an AI chatbot. These findings suggest that AI may strengthen self-regulated learning not only by supporting writing tasks but also by helping students feel more confident and engaged throughout the writing process.

Reflection and self-assessment also emerged as important aspects of AI-supported learning. Dazzeo (2024) found that AI-assisted self-assessment encouraged students to reflect

more critically on their writing, revise their work more purposefully, and evaluate their performance more effectively. Wong and Viberg (2024) reported similar results, showing that ChatGPT supported SRL in both reading and writing by providing timely assistance and increasing learner confidence. However, they also observed that AI-generated feedback often lacked the analytical depth of teacher feedback. Their findings suggest that the value of AI feedback depends largely on whether students evaluate and use it critically rather than accepting it without reflection.

Evidence from Wang et al. (2024) reinforces this point. Most students used ChatGPT for brainstorming, developing ideas, and revising their writing, and more than 92% viewed their experience positively. However, while students generally managed the writing process effectively, they were less consistent in evaluating the quality of AI-generated responses. Some learners questioned, verified, and revised AI suggestions before using them, whereas others accepted them with little evaluation. This indicates that managing the writing process does not necessarily mean that students also monitor AI-generated information critically.

These findings are consistent with concerns raised by Apriani et al. (2024), who cautioned that excessive or uncritical reliance on AI may weaken independent thinking, creativity, and reflective learning. Taken together, the reviewed studies suggest that AI is most effective in supporting SRL when learners actively evaluate and regulate their use of AI, rather than relying on it without reflection.

Overall, the available evidence shows that AI can strengthen self-regulated learning by supporting motivation, confidence, self-monitoring, self-assessment, and reflection. However, these benefits depend not only on the capabilities of AI but also on learners' active engagement in regulating their own learning throughout the writing process.

3.2. Students' interaction with AI-generated feedback (RQ2)

A second theme emerging from the reviewed studies concerns how students engage with AI-generated feedback. Rather than accepting AI feedback passively, students were found to evaluate, interpret, question, and sometimes reject the suggestions they received (Hapsari et al., 2023; Yang et al., 2024; Zhan & Yan, 2025). Their responses were influenced by several factors, including trust in AI, language proficiency, and perceptions of feedback quality (Hapsari et al., 2023; Yang et al., 2024; Thompson, 2025).

Trust was identified as an important factor shaping students' use of AI feedback. Hapsari et al. (2023) found that students' engagement with AI-supported self-assessment depended on both motivation and trust, although this relationship differed according to writing proficiency. Higher-proficiency students tended to compare AI feedback with their instructors' comments before making revisions, whereas lower-proficiency students were more likely to rely directly on AI-generated scores. These findings suggest that students with different levels of writing proficiency engage with AI feedback in different ways.

The reviewed studies also show that students responded differently depending on the type of feedback they received. Yang et al. (2024), in their study of the Pigai automated writing evaluation system, found that most AI feedback focused on grammar, vocabulary, punctuation, and other language-related issues, while relatively little attention was given to ideas or argument development. Students generally accepted direct error corrections but were more cautious when responding to linguistic suggestions. Their engagement also changed across multiple rounds of revision, moving from simple error correction to more self-regulated revision. Because some feedback was inconsistent, students often had to decide whether to accept, modify, or reject the AI's suggestions.

A similar pattern was reported by Zhan and Yan (2025). Students generally trusted AI feedback on grammar and vocabulary but were less confident in feedback related to content and argumentation. At the same time, many students copied ChatGPT's suggestions directly instead of rewriting them in their own words. Based on these findings, the authors highlighted three competencies that support effective use of AI feedback: evaluative judgment to identify inaccurate suggestions, questioning skills to improve AI-generated responses, and reflective calibration to maintain an appropriate balance between trust and skepticism.

The reviewed studies also suggest that students' trust in AI may be influenced by more than one factor. Hapsari et al. (2023) emphasized differences in learners' writing proficiency, whereas Zhan and Yan (2025) showed that students' trust also varied according to the type of feedback they received. Together, these findings indicate that students' engagement with AI feedback is shaped by both learner characteristics and feedback characteristics.

The nature of the feedback itself was also found to influence students' responses. Thompson (2025) reported that personalized AI feedback was associated with greater writing confidence and better writing performance than generic feedback or no feedback because students perceived it as more specific and easier to apply. However, the study also found that highly detailed feedback could reduce learner autonomy when students lacked the ability to evaluate it critically. Based on these findings, Thompson (2025) suggested that AI could provide language-level and technical feedback, while instructors focus on higher-order aspects of writing, such as tone and rhetorical organization.

Overall, the reviewed evidence suggests that students engage with AI-generated feedback in different ways rather than responding to it uniformly. Some learners critically evaluate and adapt AI suggestions before incorporating them into their writing, whereas others accept or copy them with little reflection. These differences highlight the importance of students' own engagement in determining how AI feedback is used during the writing process.

4. Discussion

Taken together, the reviewed studies suggest that the relationship between AI and students' learning in academic writing is more complex than simply being beneficial or harmful. Across both research questions, AI was found to support self-regulated learning in some situations while creating challenges in others, depending on how students used and responded to it (Apriani et al., 2024; Wang et al., 2024). These findings highlight the importance of examining students' interaction with AI, rather than focusing only on the technology itself.

For RQ1, the findings support Chang et al.'s (2023) view that AI can serve as a learning partner by providing personalized guidance that encourages students to regulate their own learning. The improvements in motivation, confidence, and self-regulation reported by Wei (2023) and Apriani et al. (2024) further suggest that AI contributes not only to writing performance but also to learning processes such as goal setting, monitoring, and reflection. At the same time, the reviewed studies consistently show that AI does not replace learners' role in regulating their own learning. Improvements in SRL were observed when students actively evaluated and used AI-generated support rather than relying on it without reflection. This is consistent with Apriani et al.'s (2024) finding that excessive reliance on AI may reduce independent thinking and creativity.

The findings for RQ2 provide further insight into how students engage with AI-generated feedback. Across the studies by Hapsari et al. (2023), Yang et al. (2024), and Zhan and Yan (2025), students were not simply passive recipients of AI feedback. Instead, they

evaluated, interpreted, questioned, and sometimes rejected the suggestions they received. Yang et al. (2024) also found that students accepted direct error corrections more readily than feedback that required interpretation, such as suggestions related to language use or rhetorical organization. This suggests that different types of AI feedback may encourage different levels of learner engagement during the revision process.

Another consistent finding concerns the role of trust in shaping students' responses to AI feedback. Hapsari et al. (2023) found that higher-proficiency students were more likely to compare AI feedback with their own judgment or their instructors' comments, whereas lower-proficiency students relied more directly on AI-generated suggestions. Zhan and Yan (2025) further showed that students' trust varied according to the type of feedback they received. They generally trusted AI feedback on grammar and vocabulary but were more cautious about feedback related to content and argumentation. Together, these findings suggest that students' engagement with AI feedback is influenced by both learner characteristics and the nature of the feedback itself.

4.1 Pedagogical implications

The findings of this review suggest several implications for integrating AI into academic writing instruction.

First, AI should be used to complement, rather than replace, teachers' feedback. Thompson (2025) suggested that AI is well suited to providing language-level feedback, such as grammar, vocabulary, and mechanics, while instructors continue to guide students on higher-order aspects of writing, including content development, argumentation, and rhetorical organization. This recommendation aligns with the findings of Zhan and Yan (2025), who reported that students generally trusted AI feedback on language-related issues but were more cautious about feedback on content and argumentation.

Second, writing instruction should place greater emphasis on helping students develop feedback literacy. Zhan and Yan (2025) identified three competencies that support effective use of AI feedback: evaluative judgment, questioning skills, and the ability to balance trust with critical reflection. Rather than assuming these skills develop naturally through repeated AI use, instructors can create learning activities that encourage students to evaluate, question, and justify their use of AI-generated feedback during the writing process.

Third, students should be encouraged to verify AI-generated feedback before incorporating it into their writing. Wang et al. (2024) found that although students generally managed the writing process effectively, they were less consistent in evaluating the quality of AI-generated suggestions. Providing opportunities for students to explain or justify how they use AI feedback may therefore help strengthen self-monitoring and reflective learning.

Finally, AI should be integrated into writing instruction in ways that encourage students to remain actively involved throughout the writing process. As highlighted by Apriani et al. (2024), Marzuki et al. (2023), and Shi et al. (2025), excessive reliance on AI may weaken independent thinking and creativity. AI should therefore be used as a tool to support learners' thinking and revision rather than replace their own decision-making and problem-solving.

Overall, the findings indicate that the educational value of AI depends not only on the technology itself but also on how it is incorporated into teaching and how students engage with it. Effective classroom practices should therefore focus on promoting active learner involvement, critical evaluation, and responsible use of AI throughout the writing process.

4.2 Theoretical contribution

Beyond its practical implications, this review contributes to the growing literature on SRL in AI-supported learning environments. The findings suggest that, in line with Zimmerman's (1986) conceptualization of SRL, learners continue to play an active role in planning, monitoring, and evaluating their learning even when AI is involved. Rather than replacing these self-regulatory processes, AI changes the context in which they are carried out by introducing AI-generated suggestions that learners need to evaluate, interpret, and decide whether to use (Chang et al., 2023; Wang et al., 2024; Zhan & Yan, 2025).

The review also highlights the importance of distinguishing AI dependency from general technology use. Across the reviewed studies, AI supported learning when students used it critically and remained actively involved in the writing process, but its benefits became more limited when learners relied on AI without sufficient evaluation or reflection (Apriani et al., 2024; Wang et al., 2024). These findings suggest that AI dependency may be a useful concept for understanding how AI influences SRL and deserves further theoretical and empirical investigation.

4.3 Gaps in the existing literature

The review identified several gaps that provide directions for future research. First, although many studies have reported the positive effects of AI on writing performance, motivation, and self-regulated learning, relatively few have examined AI dependency as a distinct focus of investigation. In addition, existing studies rarely explore how AI dependency may influence different components of SRL, such as goal setting, self-monitoring, and self-evaluation.

Second, the current evidence is drawn mainly from countries and regions where AI has been widely adopted in education, including China, Hong Kong, Malaysia, and several Western contexts (Wei, 2023; Wang et al., 2024; Wong & Viberg, 2024; Yang et al., 2024; Zhan & Yan, 2025; Ozdere, 2025; Chan & Lee, 2025; Thompson, 2025). No empirical study included in this review was conducted in Vietnam, despite the growing use of AI in Vietnamese higher education. Differences in educational context, English language proficiency, and technology use suggest that findings from existing studies may not be directly transferable to Vietnamese learners.

Third, most of the reviewed studies examined AI as a learning tool or a source of feedback rather than exploring how AI may influence the processes of self-regulated learning. Although several studies showed that AI can both support and hinder SRL depending on how learners use it (Chang et al., 2023; Apriani et al., 2024; Wang et al., 2024), there is still limited understanding of how AI dependency relates to different aspects of self-regulation. Developing clearer conceptual frameworks to explain this relationship would help strengthen future research in this area.

4.4 Limitations

This review has several limitations that should be considered when interpreting the findings. Although the numbers of identified and included studies are reported, the screening and full-text eligibility assessment were conducted in a single stage rather than documented as separate steps. As a result, the review provides less procedural detail than a fully protocol-driven systematic review. In addition, no standardized quality appraisal tool was applied to assess the methodological quality of the included studies. The search was also limited to English-language publications published between 2023 and 2025, meaning that relevant studies published in other languages, including Vietnamese, or before this period may not have been captured.

Another limitation relates to the evidence base itself. Most of the reviewed studies were qualitative, small in scale, or based on short-term interventions. As a result, there is still limited evidence on the long-term relationship between AI use, AI dependency, and the development of self-regulated learning. The review also found that existing studies have not yet provided a clear conceptual distinction between different forms of AI dependency, highlighting the need for further theoretical and empirical research in this area.

In summary, the findings suggest that the educational value of AI extends beyond improving writing performance or increasing efficiency. Across the reviewed studies, AI was most beneficial when learners remained actively involved in planning, monitoring, evaluating, and reflecting on their own learning. Future research should therefore move beyond examining whether AI improves writing outcomes and investigate how AI influences students' self-regulated learning, feedback literacy, and learner autonomy over time.

5. Conclusion

This thematic literature review examined recent empirical studies on the relationship between AI, SRL, and students' interaction with AI-generated feedback in academic writing. Overall, the evidence indicates that AI can make a positive contribution to the writing process by supporting learners' motivation, confidence, self-monitoring, self-assessment, and reflective learning (Chang et al., 2023; Wei, 2023; Apriani et al., 2024; Dazzeo, 2024; Wang et al., 2024). Rather than serving only as a writing assistant, AI can also support key aspects of SRL when learners use it thoughtfully and remain actively involved in their own learning.

The review further shows that students do not simply accept AI-generated feedback without question. Instead, they evaluate, interpret, verify, and sometimes reject AI suggestions depending on their writing proficiency, trust in AI, and perceptions of feedback quality (Hapsari et al., 2023; Yang et al., 2024; Zhan & Yan, 2025). These findings highlight that feedback literacy, critical evaluation, and reflective engagement remain essential, even when AI is available to support the writing process.

At the same time, the reviewed studies consistently caution against excessive reliance on AI. Although AI can support learning in many ways, relying on it uncritically may reduce opportunities for independent thinking, reflection, and learner autonomy (Apriani et al., 2024). In other words, the educational value of AI depends not only on what the technology can do but also on how learners use it. Students who actively evaluate and regulate their use of AI are more likely to benefit from it than those who depend on it without reflection.

This review also points to several directions for future research. More studies are needed to better understand how AI dependency relates to different dimensions of SRL, how this relationship develops over time, and how it varies across educational contexts. In particular, evidence from Vietnamese higher education remains limited despite the growing use of AI, highlighting the need for research that reflects the experiences of Vietnamese EFL learners.

In conclusion, AI has considerable potential to support academic writing and SRL, but it should be viewed as a tool that complements rather than replaces learners' own thinking and decision-making. The greatest educational benefits are likely to be achieved when AI is used

to encourage reflection, informed judgment, and active engagement throughout the writing process.

References

Apriani, E., Cardoso, L., Obaid, A. J., Muthmainnah, Wijayanti, E., Esmianti, F., & Supardan, D. (2024). Impact of AI-Powered ChatBots on EFL Students' Writing Skills, Self-Efficacy, and Self-Regulation: A Mixed-Methods Study. *Global Educational Research Review*, *1*(2), 57–72. https://doi.org/10.71380/GERR-08-2024-8

Chan, C. K. Y., & Lee, K. K. W. (2025). The balancing act between AI and authenticity in assessment: A case study of secondary school students' use of GenAI in reflective writing. *Computers & Education*, *238*, 105399. https://doi.org/10.1016/j.compedu.2025.105399

Chang, D. H., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. (2023). Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability*, *15*(17), 12921. https://doi.org/10.3390/su151712921

Dazzeo, R. (2024). AI-Enhanced Writing Self- Assessment: Empowering Student Revision with AI Tools. *Journal of Technology-Integrated Lessons and Teaching*, *3*(2), 80–85. https://doi.org/10.13001/jtilt.v3i2.9119

Gao, J., Zhang, J., & Li, Y. (2025). Do AI chatbot-integrated writing tasks influence writing self-efficacy and critical thinking ability? An exploratory study. *Computers and Education: Artificial Intelligence*, *9*, 100472. https://doi.org/10.1016/j.caeai.2025.100472

Hapsari, Y. T., Drajati, N. A., & Setyaningsih, E. (2023). University students' cyclical self-assessment process mediated by artificial intelligence in academic writing. *Journal of English Language Teaching Innovations and Materials (Jeltim)*, *5*(1), 31. https://doi.org/10.26418/jeltim.v5i1.59858

Hornberger, M., Bewersdorff, A., Schiff, D. S., & Nerdel, C. (2025). A multinational assessment of AI literacy among university students in Germany, the UK, and the US. *Computers in Human Behavior: Artificial Humans*, *4*, 100132. https://doi.org/10.1016/j.chbah.2025.100132

Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics*, *1*, 102324. https://doi.org/10.29140/dal.v1.102324

Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Ozdere, M. (2025). AI in Academic Writing: Assessing the Effectiveness, Grading Consistency, and Student Perspectives of ChatGPT and You.com for EFL Students. *International Journal of Technology in Education*, *8*(1), 123–154. https://doi.org/10.46328/ijte.1001

Shi, H., Chai, C. S., Zhou, S., & Aubrey, S. (2025). Comparing the effects of ChatGPT and automated writing evaluation on students' writing and ideal L2 writing self. *Computer Assisted Language Learning*, 1–28. https://doi.org/10.1080/09588221.2025.2454541

Shibani, A., Knight, S., Kitto, K., Karunanayake, A., & Buckingham Shum, S. (2024). Untangling Critical Interaction with AI in Students' Written Assessment. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3613905.3651083

Thompson, S. (2025). AI-Driven Personalized Feedback: Impacts on Undergraduates' Writing Self-Efficacy and Writing Performance. *Psychology of Education and Learning Sciences*|, *1*(1), 51–62.

Wang, C., Li, Z., & Bonk, C. (2024). Understanding self-directed learning in AI-Assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence*, *6*, 100247. https://doi.org/10.1016/j.caeai.2024.100247

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, *14*, 1261955. https://doi.org/10.3389/fpsyg.2023.1261955

Wong, J., & Viberg, O. (2024). *Supporting Self-Regulated Learning with Generative AI: A Case of Two Empirical Studies*. *LAK 2024*, 223–229.

Yang, H., Gao, C., & Shen, H. (2024). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, *29*(4), 3837–3858. https://doi.org/10.1007/s10639-023-11991-3

Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1–14. https://doi.org/10.1080/02602938.2025.2471821

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5

#155 FROM REPRODUCTION TO INTERACTION: EXTENDING ESP LISTENING PRACTICES TOWARD COMMUNICATIVE LISTENING IN AN ESP COURSE AT THE DIPLOMATIC ACADEMY OF VIETNAM (DAV)

Giang Hoang Thi Linh[1]

Abstract

Vietnam's ongoing shift toward English as an additional language for international communication has placed new demands on university English courses, particularly in English for Specific Purposes (ESP) programmes where listening instruction has traditionally emphasised comprehension and reproduction rather than communicative interaction. This pattern is especially evident at the Diplomatic Academy of Vietnam (DAV), where students engage with authentic news content on international relations, diplomacy, and economics but are rarely required to respond to it spontaneously.

This study investigated whether a brief, interaction-oriented post-listening activity could strengthen students' listening-to-respond performance, defined as the capacity to process spoken input and produce a relevant oral response. A four-week classroom-based action research project was conducted with 30 third-year undergraduate students in one ESP II listening class at DAV. The existing reproduction-based lesson sequence was left unchanged; a short peer-questioning stage was added immediately afterward.

Pre-test and post-test recordings were scored on a four-category rubric assessing comprehension, response relevance, fluency, and response development. Because pre-post difference scores departed from normality, Wilcoxon signed-rank tests served as the primary inferential analysis, supported by paired-samples *t*-tests as a robustness check. Both tests indicated statistically significant gains across all categories and the total score (all $p < .01$), with effect sizes ranging from medium-to-large ($d = 0.66$) to very large ($d = 2.10$), the strongest gain occurring in response development. Qualitative analysis across multiple news genres corroborated this pattern, showing a consistent shift from factual recall toward explanation and elaboration. The findings suggest that minimal, curriculum-compatible interaction extensions can meaningfully advance communicative listening within reproduction-oriented ESP instruction.

***Keywords*:** ESP listening, listening-to-respond, interaction, Vietnam, communicative competence, action research

1. Introduction

Vietnam's deepening international integration has raised expectations for communicative English competence across higher education, requiring university English programmes to prepare students not merely to know the language but to use it in real academic and professional exchange. Following Canale and Swain (1980), communicative competence denotes the ability to use language appropriately and effectively, grammatically, sociolinguistically, discursively, and strategically, across authentic contexts, a standard

[1] *Lecturer, English Faculty, Diplomatic Academy of Vietnam, Hanoi, Vietnam.*

increasingly written into Vietnamese higher-education policy even where classroom practice has been slower to follow.

English for Specific Purposes (ESP) courses sit at the centre of this shift because they exist specifically to prepare learners for discipline-specific communication. At DAV, ESP listening courses draw on authentic, news-based material tied to international relations, diplomacy, economics, and international law, content that mirrors real diplomatic and legal communication. Students are streamed by proficiency at entry; those without an IELTS score of 6.0 or above sit an institutional placement test aligned with the Vietnamese six-level foreign-language framework, itself broadly mapped to the CEFR. Students in ESP II, the level examined here, correspond approximately to B1.

In practice, instruction at this level centres on comprehension and reproduction: students listen to authentic news and reproduce its content, in writing or speech, as accurately as possible. This builds comprehension, vocabulary, and information-processing skill, but leaves little room for learners to respond to what they hear, to interpret it, react to it, or use it communicatively. Contemporary TESOL scholarship increasingly foregrounds listening-to-respond, the capacity to process spoken input and produce a meaningful spoken reaction to it, yet bringing this into heavily structured ESP settings is not straightforward given curriculum load, assessment expectations, and learners' proficiency ceilings.

The problem motivating this study arose from the researcher's own classroom observation in ESP listening courses at DAV. Students could generally identify key information in authentic news reports, but many struggled to respond to it spontaneously: responses were often short, memorised, or lifted almost verbatim from the listening text, and students hesitated when asked to elaborate or explain cause and effect. This reflects a structural feature of the curriculum rather than an individual shortcoming, since comprehension and reproduction are heavily practised and assessed while spontaneous, meaning-based response is neither taught nor tested. A full communicative redesign was not a realistic option within the existing course structure, so the problem this study addresses is narrower: whether a small extension to existing reproduction tasks can move students toward more communicative listening behaviour without disturbing the curriculum they are already required to deliver.

Accordingly, the study pursues three aims: to describe a brief, interaction-oriented post-listening activity compatible with an existing ESP reproduction-based sequence; to examine changes in listening-to-respond performance, comprehension, response relevance, fluency, and response development, across a four-week implementation; and to interpret these changes, together with classroom observation and recorded excerpts, in relation to theory and research on interaction, output, and ESP pedagogy. The study addresses the following research question:

To what extent can brief interaction-oriented post-listening activities enhance students' listening-to-respond performance in an ESP listening course?

This question is framed as a matter of extent rather than a directional prediction because the study is a single-cycle, single-group action research project conducted by the class teacher within one intact class over four weeks. Under these conditions, no control group, no random assignment, teacher and researcher combined in one role, a research question is a more defensible instrument than a formal hypothesis, which would imply a confirmatory logic the design was not built to support (Burns, 2010; Mills, 2011).

2. Literature Review

2.1. ESP Listening in Higher Education

Listening sits at the centre of ESP instruction because learners must process discipline-specific input in academic and professional settings. Dudley-Evans and St John (1998) argue that ESP instruction should be built around learners' communicative needs within a specific domain, typically lectures, interviews, presentations, and authentic media. Authentic listening material builds familiarity with real-world language use and professional communication patterns (Gilmore, 2007), though the same authenticity raises the cognitive load learners must manage. Several scholars have pushed ESP listening instruction to move past comprehension testing toward communicative application: Flowerdew and Miller (2005) observe that listening instruction still prioritises information recall over interaction, and Field (2008) describes listening classrooms that position learners as passive receivers. In Vietnam and across Asian EFL contexts more broadly, listening instruction has historically emphasised test preparation and factual comprehension (Nguyen & Boers, 2019), a pattern sharpened in ESP settings where anxiety about specialised content can push students toward memorisation over communicative risk-taking.

2.2. Listening-to-Respond and Communicative Listening

Older listening pedagogy treated listening largely as decoding followed by comprehension checks; communicative approaches instead treat listening as interpretive and interactive. Richards (2008) argues listening cannot be cleanly separated from speaking, since real communication requires listeners to act on what they hear, and Nation and Newton (2009) make a parallel case for teaching listening and speaking together through communicative tasks. Vandergrift and Goh (2012) push this further, arguing that effective listening instruction should cultivate active meaning processing and metacognitive reflection rather than comprehension testing alone; Vandergrift (2007) frames listening research in similar terms, as increasingly concerned with active meaning construction. Rost (2011) notes that response-oriented tasks push learners to process input more deeply, since they must prepare a reaction rather than select a correct answer, and Newton and Nation (2020) report that post-listening interaction can support both fluency and communicative confidence. This matters especially in ESP, where professional listening is rarely passive: diplomatic and international settings routinely require people to react, question, and explain, so listening-to-respond arguably reflects the target communicative demand more directly than comprehension testing does.

2.3. Interaction-Oriented Post-Listening Activities

Post-listening stages give learners room to consolidate comprehension and extend what they have heard into productive language use (Underwood, 1989). Two strands of theory motivate why peer interaction specifically should help. Long's (1996) Interaction Hypothesis holds that interaction supports acquisition because learners negotiate meaning and receive opportunities for modified output; Swain (2005) argues, complementarily, that producing language pushes learners to notice gaps in what they can already say, building communicative control. Peer interaction may also lower speaking anxiety, since learners are often more comfortable experimenting and reformulating with classmates than performing directly for the teacher. None of this is straightforward where courses remain assessment-driven: classroom time is limited and students carry real anxiety about spontaneous speaking, which is precisely why a minimal, low-preparation adaptation, rather than a larger communicative overhaul, is worth testing in a setting like this one.

2.4. Theoretical Framework

The theoretical framework guiding this study follows a single explanatory chain rather than a list of separate theories. Communicative Competence (Canale & Swain, 1980) supplies the target construct: the goal of instruction is appropriate, effective language use across grammatical, sociolinguistic, discourse, and strategic dimensions, not comprehension accuracy alone. Long's (1996) Interaction Hypothesis supplies the mechanism for why peer questioning should help learners move toward that target, since negotiating meaning with a partner creates opportunities for modified input and output that a solitary reproduction task does not. Swain's (2005) Output Hypothesis supplies the complementary production-side mechanism: producing a spoken response, rather than merely comprehending, pushes learners to notice gaps between what they want to say and what they can currently say. Listening-to-respond operationalises this chain for classroom measurement, sitting at the intersection of Long's interactional mechanism and Swain's output mechanism, and ESP pedagogy (Dudley-Evans & St John, 1998) supplies the professional-relevance rationale that justifies building the intervention around DAV's own news-based content.

2.5. Research Gap and Contribution of the Present Study

ESP listening research documents that reproduction-oriented instruction dominates ESP classrooms, including in Vietnam, and explains why this persists institutionally (Dudley-Evans & St John, 1998; Gilmore, 2007; Flowerdew & Miller, 2005; Field, 2008; Nguyen & Boers, 2019). Listening-to-respond and interaction research establishes, largely in general EFL/ESL contexts, why moving listening instruction toward interaction should support learners (Richards, 2008; Nation & Newton, 2009; Vandergrift, 2007; Vandergrift & Goh, 2012; Rost, 2011; Newton & Nation, 2020; Long, 1996; Swain, 2005). What remains unexplored is the middle ground: whether a minimal, curriculum-embedded interaction stage that leaves existing reproduction tasks untouched can produce measurable movement toward listening-to-respond inside an ESP course structured around reproduction, in a context where wholesale communicative redesign is not realistic. The present study's contribution is this bridging step, tested statistically rather than described only in general terms.

3. Methodology

3.1. Research Design

The study used a classroom-based action research design, structured as a one-group pre-test/post-test comparison over a four-week intervention period. Action research was appropriate because the aim was to improve classroom practice within a real instructional context, testing whether a communicative extension could be added to an existing ESP curriculum under authentic institutional constraints, rather than to establish a causal effect under controlled experimental conditions (Burns, 2010; Mills, 2011). The single ESP II class involved was selected through convenience sampling: it was the class the researcher was already teaching that semester, which made a genuine classroom-embedded trial possible. Other ESP classes taught by the same researcher continued with regular instruction but did not function as a formal control group, since students were not randomly assigned and no parallel data were collected from them, a design limitation discussed further below.

3.2. Participants

Participants were 30 third-year undergraduate students in one ESP II listening class at DAV, placed at approximately B1 level under DAV's institutional procedures. Students without an internationally recognised certificate sat an institutional placement examination

covering listening, grammar, vocabulary, reading, and short written responses, aligned with the Vietnamese six-level framework. Most participants had studied English through secondary school before entering university, and several had attended specialised or high-performing secondary schools, consistent with DAV's competitive admissions profile. The course met once weekly for 135-minute sessions; before the intervention, instruction had been predominantly comprehension-oriented, with limited communicative follow-up.

The researcher was also the classroom teacher, a dual role that carries an inherent risk of bias, since the same person designed the intervention, taught it, and scored the outcome measure. Before data collection, students were informed of the study's purpose and gave verbal consent for anonymised use of their recordings; participation did not affect course grades. Ethical approval was granted by the faculty research committee of the English Division at DAV.

3.3. Instructional Context

The course covered listening material on international relations, economics, diplomacy, international law, and global affairs, using authentic sources such as BBC World News, CNN International, Al Jazeera English, and VOA Learning English. Standard practice involved listening comprehension exercises, vocabulary development, note-taking, and news reproduction, with the emphasis on comprehension and recall rather than spoken production. During the intervention, listening texts covered international trade, climate policy, regional conflict, and diplomatic relations, chosen because they matched the curriculum's existing thematic focus. Each 2.5-hour lesson followed the textbook's usual sequence, pre-listening preparation, listening comprehension, vocabulary consolidation, note-taking, and news reproduction, and the intervention was built to sit inside this structure rather than compress or replace any part of it.

3.4. Intervention

The intervention added a brief, interaction-oriented post-listening stage immediately after the standard reproduction task, without replacing any existing activity. A typical lesson followed this sequence: warm-up and vocabulary preparation; authentic news listening and comprehension tasks; vocabulary consolidation and note-taking; oral or written news reproduction; pair questioning and listening-to-respond activity; and teacher summary. After reproduction, students worked in pairs, one answering follow-up questions while the partner asked, using teacher-provided prompts; roles rotated so both students practised questioning and responding. Prompts moved from factual clarification in the first two weeks to more elaborated, causal-reasoning prompts in the final two weeks (e.g., What happened? Why is it important? How are people reacting? What may happen next?), matched to text structure. The teacher circulated, took brief observation notes, and intervened only occasionally, keeping corrective feedback deliberately light so as not to suppress spontaneous response. The interaction stage lasted roughly three to seven minutes per lesson.

Table 1

Sample Lesson Structure During the Intervention

Stage	Activity	Approximate Time
Warm-up	Topic discussion and activation of background knowledge	15 minutes
Pre-listening	Vocabulary preview and prediction activities	25 minutes

Listening	Authentic news listening and comprehension tasks	40 minutes
Reproduction	Oral or written news reproduction	25 minutes
Interaction extension	Pair questioning and listening-to-respond activity	3–7 minutes
Wrap-up	Teacher summary and clarification	10 minutes

3.5. Data Collection

Pre-test and post-test listening-to-respond tasks were administered before and after the four-week intervention. Students listened to a short news report, took brief notes, and recorded oral responses to two follow-up questions, without preparing full scripts in advance; each response lasted roughly 45–60 seconds. Recordings were scored by the researcher using the rubric described below. A second independent rater was not involved, so inter-rater reliability could not be established, a limitation addressed in the Discussion. The researcher also kept brief observation notes on participation patterns, hesitation, and interactional difficulties during the pair activities, used to support interpretation of the qualitative findings.

3.6. Assessment Rubric

Recordings were scored on a four-category analytic rubric, Comprehension, Response Relevance, Fluency, and Response Development, each rated on a five-point scale (Table 2).

Table 2

Summary of Analytic Rubric Descriptors

Category	Score 1	Score 3	Score 5
Comprehension	Limited understanding of main ideas	Understands main ideas with some missing details	Accurate understanding of main ideas and supporting details
Response Relevance	Responses minimally connected to questions	Responses generally relevant but occasionally unclear	Responses consistently relevant and appropriate
Fluency	Frequent pauses and fragmented speech	Some hesitation but generally understandable	Smooth and sustained speech with minimal hesitation
Response Development	Very short or repetitive responses	Some explanation or elaboration provided	Responses extended with clear explanation and supporting ideas

The rubric was developed by the researcher directly from the constructs central to the study's theoretical framework, since no published instrument targeted listening-to-respond specifically for an ESP setting. Comprehension descriptors reflect established listening-assessment criteria describing accuracy of understanding of main ideas and supporting detail (Field, 2008; Flowerdew & Miller, 2005). Response Relevance and Response Development descriptors were written to capture the interactional and output dimensions emphasised by Long (1996) and Swain (2005): whether a response addressed the question asked, and whether it was extended with explanation rather than left at bare recall. Fluency descriptors reflect the broader communicative competence construct (Canale & Swain, 1980), specifically its emphasis on effective, connected spoken performance rather than accuracy alone. Score-point

descriptors were written in plain, behaviourally anchored language so that a single classroom rater could apply them consistently across recordings.

Conceptual grounding of this kind, rather than statistical validation, is the basis on which the rubric's appropriateness for this study rests. No formal expert panel review, pilot validation, or psychometric analysis, such as a reliability coefficient or factor analysis, was conducted, and a single rater scored all recordings, so no inter-rater reliability statistic exists. For a small-scale classroom action research project designed to trial a pedagogical adjustment rather than to produce a standardised instrument, this level of instrument development is consistent with the exploratory purpose of the study.

3.7. Delimitations of the Study

The scope of this study was deliberately bounded, and these boundaries should be read as intentional design decisions rather than as unplanned gaps, which are discussed separately under Limitations. The study examined one intact ESP II class at one institution because the design required a genuine classroom-embedded trial, led by the students' own teacher, rather than an artificial or borrowed setting; investigating multiple classes was beyond the scope of a single-cycle action research project of this kind. The study examined listening-to-respond performance specifically, rather than general oral proficiency or listening comprehension in isolation, because this construct most directly reflects the interactional demand the intervention was designed to strengthen. Only authentic news-based listening lessons already used in the ESP II syllabus were included, so that the intervention could be evaluated under the exact instructional conditions the course normally provides, rather than under specially constructed materials that might not transfer to regular practice. Finally, the intervention was limited to a brief interaction-oriented post-listening questioning stage rather than a broader communicative task, such as role play or extended discussion, because the guiding concern was whether a minimal, low-preparation adaptation could be absorbed into an existing reproduction-based lesson without displacing it; broader tasks would have required curriculum time this study was not designed to test.

3.8. Data Analysis

Scores for each rubric category were entered into Microsoft Excel for both the pre-test and post-test, and descriptive statistics, means, standard deviations, and mean differences, were computed for each category and the total score. Prior to inferential testing, the assumption of normality was checked by applying the Shapiro-Wilk test to the pre-post difference scores for each category and the total score, since paired-samples *t*-tests assume that these difference scores are approximately normally distributed. In every case, the Shapiro-Wilk test indicated a significant departure from normality, W ranging from 0.58 to 0.84, $p < .001$, consistent with the bounded, discrete nature of a five-point rubric applied to a sample of this size.

Because this assumption was not satisfied, the non-parametric Wilcoxon signed-rank test served as the primary inferential analysis comparing pre-test and post-test scores for each category and the total score. Paired-samples *t*-tests were also computed and are reported alongside the Wilcoxon results as a supplementary robustness check; because both tests converged on the same conclusion for every comparison, this convergence increases confidence that the pattern of significant gains is not an artefact of the specific test chosen. Cohen's d for paired samples (the mean difference divided by the standard deviation of the difference scores) was computed for each comparison as a standardised, scale-independent measure of the size of the change, interpreted using Cohen's (1988) conventional benchmarks (0.2 = small, 0.5 = medium, 0.8 = large). All statistics were computed using JASP.

Qualitative analysis involved examining representative excerpts from student recordings and classroom observation notes, organised by recurring news-text types and question types used throughout the intervention, to identify patterns in explanation, causal reasoning, elaboration, and response relevance.

4. Results

4.1. Descriptive Statistics

All 30 students completed both the pre-test and post-test tasks. Table 3 presents the mean scores, standard deviations, and mean differences for each rubric category and the total score.

Table 3

Descriptive Statistics for Pre-Test and Post-Test Scores (N = 30)

Category	Pre-test M (SD)	Post-test M (SD)	Mean Difference
Comprehension	3.37 (0.61)	3.70 (0.60)	0.33
Response Relevance	3.03 (0.67)	3.47 (0.63)	0.43
Fluency	2.87 (0.68)	3.23 (0.68)	0.37
Response Development	2.53 (0.68)	3.23 (0.63)	0.70
Total Score (/20)	11.80 (2.38)	13.63 (2.25)	1.83

Scores improved across all four categories, with the largest raw gain in Response Development (0.70), followed by Response Relevance (0.43), Fluency (0.37), and Comprehension (0.33), consistent with an intervention targeting interaction and explanation specifically, layered onto a curriculum that had already trained comprehension and reproduction intensively.

4.2. Inferential Statistics

Shapiro-Wilk tests on the pre-post difference scores indicated a significant departure from normality for every category and the total score (Comprehension: $W = 0.60$, $p < .001$; Response Relevance: $W = 0.63$, $p < .001$; Fluency: $W = 0.72$, $p < .001$; Response Development: $W = 0.58$, $p < .001$; Total: $W = 0.84$, $p < .001$). Consistent with this violation, Wilcoxon signed-rank tests served as the primary inferential analysis (Table 4).

Table 4

Wilcoxon Signed-Rank Test Results and Effect Sizes (N = 30)

Category	z	p	r	Cohen's d
Comprehension	-2.80	.002	.89	0.70
Response Relevance	-3.18	<.001	.88	0.86
Fluency	-2.69	.003	.75	0.66
Response Development	-4.02	<.001	.88	1.50
Total Score	-4.54	<.001	.87	2.10

All five comparisons were statistically significant, indicating that the improvement from pre-test to post-test was unlikely to reflect chance variation. Effect sizes ranged from

medium-to-large for Fluency (d = 0.66) and Comprehension (d = 0.70) to large for Response Relevance (d = 0.86) and very large for Response Development (d = 1.50) and the total score (d = 2.10); the corresponding rank-biserial correlations (r = .75-.89) point to the same conclusion.

Paired-samples t-tests, reported in Table 5 as a supplementary robustness check, converged with the Wilcoxon results for every comparison, and the 95% confidence intervals for the mean difference excluded zero in every case.

Table 5

Paired-Samples t-Test Results (Supplementary Robustness Check, N = 30)

Category	t(29)	p	95% CI for MD
Comprehension	3.81	<.001	[0.15, 0.51]
Response Relevance	4.71	<.001	[0.25, 0.62]
Fluency	3.61	.001	[0.16, 0.57]
Response Development	8.23	<.001	[0.53, 0.87]
Total Score	11.49	<.001	[1.51, 2.16]

Convergence across parametric and non-parametric tests strengthens confidence that the observed gains reflect a genuine pattern in the data rather than an artefact of a particular statistical procedure. The two categories showing the largest effect sizes, Response Development (d = 1.50) and the total score (d = 2.10), indicate that the intervention's clearest influence was on students' capacity to elaborate on and explain what they had heard, rather than on comprehension accuracy alone, a distinction developed further in the Discussion.

4.3. Qualitative Findings

Three cross-cutting patterns emerged when recordings and observation notes were analysed thematically across the recurring news genres used throughout the intervention, rather than lesson by lesson. Responses moved from recall toward explanation, became more closely anchored to the partner's question than to the source text, and this shift recurred across genres with different communicative demands rather than being confined to one or two familiar topics. Each pattern is illustrated below with excerpts drawn from more than one lesson of the relevant type.

The most consistent pattern was a shift from bare recall toward causal explanation, visible in both breaking-news and policy-related lessons, such as coverage of the Ever Given's grounding in the Suez Canal and of the Regional Comprehensive Economic Partnership (RCEP). Prompts for this kind of content, including What happened?, Why is this important?, and What could happen next?, were designed to move students from reporting an event to interpreting its significance. Before the intervention, responses of this type were typically limited to bare fact: one pre-test response to the Suez Canal report stated only that “the ship blocked the canal. Trade stopped,” and a comparable pre-test response to RCEP reduced the agreement to a label, “RCEP is a trade agreement.” Post-test responses to similar content added an explanatory clause in both cases: “The ship blocked the canal, so many ships could not pass. This affected transportation and world trade,” and “RCEP is a trade agreement and it can help countries recover their economies after the pandemic.” Because this shift recurred across two unrelated genres rather than in a single lesson, it illustrates the Response Development gain

reported in Table 4 and points to a general change in how students used spoken input rather than a response tied to one topic.

A second pattern was closer alignment between responses and the partner's specific question rather than the original listening text, evident in both economic-news and organisational-decision lessons, such as coverage of sanctions affecting Russian seafood exports and of a Volkswagen labour dispute. Prompts of this kind, including How did this affect different people or countries?, Who benefited or was disadvantaged?, and Do you agree with the decision? Why?, asked students to reason about consequence or to take a position rather than restate content. Early responses to both genres tended to restate the reported fact without engaging the interactional question, reflecting adequate Comprehension but weaker Response Relevance. By the later weeks, students more often addressed the partner's follow-up directly, for instance distinguishing which countries' exporters were affected in the sanctions lessons, and linking a company's decision to its effect on workers before offering a judgement in the labour-dispute lessons. This convergence across an explanatory and an evaluative genre is broadly consistent with the negotiation-of-meaning mechanism proposed by Long (1996), in which the partner's question, rather than the source text, increasingly shapes what a learner says next.

A third pattern, cutting across the first two, was that neither shift appeared confined to a single genre or question type. Explanatory gains were observed in both factual (breaking news) and policy-oriented (international agreements) content, and relevance gains were observed in both economic and evaluative (organisational-decision) content. Taken together, this suggests a change in listening-to-respond behaviour that extended across the kinds of authentic material the course already used, rather than an improvement specific to one or two familiar lessons.

Considered together, this pattern is broadly consistent with the quantitative gains reported above: the largest effect sizes were observed for Response Development and Response Relevance, the two categories most directly reflected in the qualitative shifts described here. Given the study's design, this convergence does not by itself establish that the intervention caused the change, but it does support a fuller reading of the results than either the statistical or the qualitative evidence would offer alone, since the shift toward explanation and relevance appears to have been distributed across genres rather than isolated to particular lessons.

5. Discussion

5.1. Differential Gains Across Rubric Categories

One consistent finding across the quantitative and qualitative evidence is that the intervention shifted the character of students' listening-to-respond performance from factual reporting toward explanation, and did so across genres rather than within a single familiar lesson. This pattern is broadly consistent with Swain's (2005) Output Hypothesis: requiring a spoken response, rather than comprehension alone, is the condition under which learners are thought to notice gaps between what they understand and what they can currently articulate, and Response Development, the category most directly tied to that mechanism, produced the largest effect size observed (d = 1.50). Response Relevance improved almost as strongly (d = 0.86), a result perhaps better explained by Long's (1996) Interaction Hypothesis than by output alone, since negotiating meaning with a peer gives students real-time signals about whether a response addresses the question, a mechanism visible in the qualitative shift from restating the listening text toward answering a partner's specific follow-up.

By comparison, Comprehension showed the smallest, though still significant, effect (d = 0.70). This is not a surprising result: the existing curriculum had already trained comprehension and reproduction intensively, consistent with Nguyen and Boers's (2019) account of Vietnamese and broader Asian EFL contexts, where listening instruction has long emphasised test preparation and factual comprehension, precisely the skill this cohort had already practised. Fluency showed a similar pattern (d = 0.66), plausibly because the intervention targeted relevance and elaboration rather than fluency directly, and because four weeks is a short period in which to expect observable change in spontaneous speech at approximately B1 level. Read together, these differential effects, largest where the intervention's mechanism was most direct, offer a more precise account of what changed than the total score alone.

5.2. Extending ESP Listening Research through Minimal Interaction

One finding that deserves particular attention is that the largest improvements occurred in Response Development and Response Relevance rather than in Comprehension. In other words, students did not simply become better at understanding the listening texts; they became more willing to explain, interpret, and respond to what they had heard. This distinction helps position the study within existing ESP listening research. While previous studies have shown that reproduction activities support comprehension and information organisation (Dudley-Evans & St John, 1998; Flowerdew & Miller, 2005; Field, 2008; Long, 1996; Swain, 2005; Vandergrift & Goh, 2012), relatively few have examined whether a brief interaction stage can build on those strengths without changing the overall lesson structure.

The present findings offer preliminary support that it can: reproduction and interaction need not be framed as competing approaches to listening instruction. Reproduction continues to do what Field (2008) and Flowerdew and Miller (2005) describe it as already doing well, supporting comprehension and information organisation, while a short interaction stage appears to do complementary work, encouraging communicative use of that information, in line with Richards's (2008) argument that listening and speaking are not easily separated in instruction.

Such a complementary relationship may matter most where a full communicative redesign is hardest to justify. In assessment-driven ESP programmes, where curriculum time and assessment weightings are fixed, a three-to-seven-minute addition seems more likely to be adopted and sustained than a larger structural change would be. For ESP teachers in similar positions, the practical question this study speaks to is less whether to replace reproduction with communicative tasks than what minimal addition might create room for response, a question addressed here with one tested, low-cost example rather than a general prescription.

5.3. Relevance to Vietnam's Shift toward Communicative English

Beyond the immediate classroom, these findings speak to a broader transition underway in Vietnamese higher education, in which English instruction is gradually repositioning English less as a subject mastered for examinations and more as an additional language for academic and professional use. Vandergrift and Goh (2012) argue that this kind of transition requires listening instruction to build active meaning processing and metacognitive engagement gradually rather than in short bursts, and the present study offers one classroom-tested mechanism for incremental progress toward that goal without waiting for curriculum-wide reform. Because DAV's ESP listening courses are built on internationally oriented content, diplomacy, law, and economics, where responding to spoken input is arguably a professional requirement rather than a general communicative skill, the case for narrowing the gap with

communicative competence (Canale & Swain, 1980) may be stronger here than in general English contexts.

5.4. Limitations

Several limitations bound these findings. The study drew on a single class at a single institution, so results cannot be generalised without further testing. The absence of a control group means the gains reported are consistent with the intervention rather than proof of it. The researcher's dual role as teacher and rater introduces a structural source of bias, and reliance on a single rater means no inter-rater reliability statistic could be established. The four-week window is sufficient to detect an initial shift but not to establish durability, since no delayed post-test was administered. These limitations should be considered when interpreting the findings. At the same time, the consistency between the quantitative results and the classroom evidence provides reasonable confidence that the observed changes reflected meaningful developments in students' listening-to-respond performance.

6. Conclusion

The present study asked whether a brief, interaction-oriented post-listening activity could strengthen students' listening-to-respond performance in an ESP listening course at the Diplomatic Academy of Vietnam. Wilcoxon signed-rank tests, corroborated by paired-samples *t*-tests, showed statistically significant gains across comprehension, response relevance, fluency, and response development, with effect sizes ranging from medium-to-large to very large, and qualitative evidence across multiple news genres, breaking news, economic reporting, international agreements, and organisational decisions, pointed to the same underlying shift: from factual recall toward explanation, causal reasoning, and reasoned opinion.

Read against Canale and Swain's (1980) communicative competence framework and the interactional and output mechanisms proposed by Long (1996) and Swain (2005), this pattern suggests that reproduction-based ESP listening instruction does not need to be abandoned in order to move toward communicative goals; it can instead be extended. A brief interaction-oriented stage, added after an unchanged reproduction task, appears sufficient to convert a comprehension exercise into a modest bridge toward communicative listening. For ESP teachers working within fixed curricula and assessment structures, of the kind common across Vietnamese higher education, this offers one practical, evidence-based option that does not require institutional redesign to implement.

The broader implication is a modest one: even within reproduction-bound ESP classrooms, a few minutes of structured interaction may be enough to turn listening into a starting point for communication rather than an endpoint in itself.

Declaration of AI Use

During the preparation of this manuscript, the author used Claude and ChatGPT to improve the stylistic clarity, readability, and grammatical accuracy of the text. The underlying conceptual framework, classroom intervention design, data collection, statistical analyses, and qualitative interpretations were executed entirely by the author, who maintains full accountability for the integrity of the final content.

References

Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. Routledge.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/I.1.1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.

Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge University Press.

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*(2), 97–118. https://doi.org/10.1017/S0261444807004144

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Mills, G. E. (2011). *Action research: A guide for the teacher researcher* (4th ed.). Pearson.

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.

Newton, J. M., & Nation, I. S. P. (2020). *Teaching ESL/EFL listening and speaking* (2nd ed.). Routledge.

Nguyen, C.-D., & Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED Talk. *TESOL Quarterly, 53*(1), 5–29. https://doi.org/10.1002/tesq.441

Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.

Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Routledge.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–484). Lawrence Erlbaum Associates.

Underwood, M. (1989). *Teaching listening*. Longman.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191–210. https://doi.org/10.1017/S0261444807004338

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Bionote

Giang Hoang Thi Linh is a lecturer of English for Specific Purposes at the Diplomatic Academy of Vietnam. Her teaching and research interests include ESP pedagogy, listening instruction, communicative language teaching, and classroom-based action research in Vietnamese higher education contexts.

#156 EXPLORING THE IMPACTS OF OVER-RELIANCE ON GENERATIVE AI TOOLS IN ESSAY WRITING: A CASE STUDY OF HUFLIT SOPHOMORES

Pham Kim Quoc[1], Le Hoang Yen Vy[2], Tran Nguyen Thu Ha[3], Bui Hoang Gia Han[4], Tran Thi Truc Duyen[5]

Abstract

Artificial intelligence tools, especially Generative AI (GenAI), are increasingly used in higher education, particularly in academic writing skills. Among these tools, ChatGPT has become popular among university students because it can generate ideas, correct grammar, and help to organize essays. However, excessive reliance on ChatGPT may negatively influence students' independent thinking and writing development. Since essay writing is an essential skill for English-major students, understanding the risks of overusing AI tools has become increasingly important. Therefore, this study was conducted to investigate the impacts of utilizing GenAI tools, especially ChatGPT, in essay writing among sophomores at the University of Foreign Languages and Information Technology (HUFLIT). The study involved the participation of 127 HUFLIT sophomores who had experience using ChatGPT in essay writing. The questionnaire of a 5-point Likert for 121 participants and the semi-structured interviews for six were administered to the students, examining four aspects: dependence on AI-generated content, reduction in independent idea generation, impact on critical thinking and the overall development of writing autonomy. The results revealed that students perceived the impact of overusing ChatGPT at a moderate level. The study focuses on evaluating the potential negative impacts and the level of student dependency on writing skills; in addition, several pedagogical implications were suggested to promote responsible AI use and support students' independent writing development.

Keywords: ChatGPT, AI-assisted writing, essay writing skills, English-major sophomores

1. Introduction

1.1. Overview of the Background Information

Artificial intelligence (AI) is rapidly transforming higher education, particularly through the integration of generative AI tools into teaching and learning practices. Recent research highlights that generative AI technologies, such as ChatGPT, are increasingly embedded in academic contexts to support writing instruction, provide automated feedback, and facilitate personalized learning experiences (Ding et al., 2026; Crosthwaite & Sun, 2025).

In today's globalized world, the field of English Language Studies (ELS) has gained greater significance, as English is widely used for international communication, education, and

[1] *Pham Kim Quoc, first and corresponding author, English Language Teaching senior, Faculty of Foreign Languages, Huflit, email: 23DH713538@st.huflit.edu.vn*
[2] *Le hoang Yen Vy, English Language Teaching senior, Faculty of Foreign Languages, Huflit*
[3] *Tran Nguyen Thu Ha, English Language Teaching senior, Faculty of Foreign Languages, Huflit*
[4] *Bui Hoang Gia Han, English Language Teaching senior, Faculty of Foreign Languages, Huflit*
[5] *Tran Thi Truc Duyen, English lecturer, Faculty of Foreign Languages, Huflit*

professional purposes. English-majored students are expected to develop proficiency in four main language skills: listening, speaking, reading, and writing. Writing is an essential skill in academic learning because it helps students organize ideas, develop arguments, and communicate their thoughts clearly (Flower & Hayes, 1981; Hyland, 2003; Wingate, 2012). Beyond educational contexts, writing is also important for professional and personal communication (Weigle, 2002). To write effectively, students majoring in English need not only grammatical knowledge and vocabulary but also the ability to generate ideas, organize content logically, think critically, and develop arguments coherently (Kellogg, 2008; Setyowati et al., 2020; Wingate, 2012).

Because of these writing requirements, generative AI tools have become more relevant to English academic writing. ChatGPT is commonly used by students as a form of writing support because it can help them brainstorm ideas, plan essay structures, revise grammar, improve vocabulary, and receive immediate feedback on their drafts. Recent studies have shown that generative AI technologies are being integrated into higher education to support writing instruction, provide automated feedback, and facilitate personalized learning experiences (Crosthwaite & Sun, 2025; Ding et al., 2026). In EFL contexts, AI-assisted writing tools have also been reported to improve students' writing performance and confidence (Jmaiel et al., 2025). However, the growing use of ChatGPT in essay writing also raises concerns about students' dependence on AI-generated content and their ability to write independently.

1.2. Problem Statement

The use of ChatGPT has become common among HUFLIT sophomores in essay writing courses as it can assist them in key stages of the writing process such as generating ideas, planning essay outlines, organizing arguments, and improving word choices. When used appropriately, ChatGPT can function as a useful learning tool that supports students during the writing process; however, problems arise when students rely on ChatGPT too frequently and let it complete major parts of their essays for them. Instead of brainstorming and coming up with their own ideas and writing strategies, some students may depend on ChatGPT to provide ready-made content.

This over-reliance may reduce students' opportunities to practice essential writing processes on their own. This lack of practice overtime may make them gradually become less capable of generating original ideas, organizing arguments logically, and evaluating the quality of their own writing. As a result, their independent writing competence may decline, which may then reduce their confidence when they are required to write essays without ChatGPT support. Furthermore, by allowing ChatGPT to organize arguments and generate content automatically, students may bypass important cognitive processes involved in developing their original reasoning. Recent studies have also warned that excessive dependence on ChatGPT can negatively affect students' creativity, critical thinking ability, and long-term writing development (Zhang et al., 2024). Therefore, further investigation is needed to examine how the overuse of ChatGPT may affect HUFLIT sophomores' ability to generate ideas, develop arguments, evaluate their own writing, and complete essay-writing tasks independently.

1.3. Objectives

This study aims to investigate the potential risks associated with the overuse of generative AI tools in essay writing, with a specific focus on ChatGPT among HUFLIT sophomores. As ChatGPT is one of the most widely used generative AI tools for academic writing support, the study examines the extent to which students depend on ChatGPT-generated content during the writing process. It also explores how this dependence may influence their

ability to generate ideas, develop arguments, and write more independently. Since academic essay writing requires active thinking, continuous practice, and independent decision making, the study seeks to determine whether the overuse of ChatGPT may limit students' engagement in these essential writing processes. To achieve these objectives, the study is guided by the following research questions:

1. To what extent does overuse of ChatGPT lead to students' dependence on AI-generated content in essay writing?

2. How does the excessive use of ChatGPT influence HUFLIT sophomores' ability to think independently when developing arguments in essays?

1.4. Scope of the Study

This study is limited to English Language Studies sophomores at HUFLIT who have used ChatGPT to support their English essay writing. The study focuses on students' overuse of ChatGPT in essay-writing tasks because these tasks require a combination of cognitive and linguistic processes, including idea generation, argument development, content organization, language use, and revision. More specifically, it investigates students' dependence on ChatGPT-generated content and how this dependence may affect their ability to think independently, develop arguments, self-edit, and complete essays without excessive support from ChatGPT.

Drawing on Cognitive Load Theory, this study views academic essay writing as a cognitively demanding task in which students need to manage several processes simultaneously, such as understanding the topic, generating ideas, organizing arguments, selecting appropriate language, and revising their work. As these processes may create a high cognitive load, particularly for students who are still developing their academic writing ability, the overuse of ChatGPT may become problematic when they rely on the tool to handle demanding stages of essay writing rather than actively practicing the thinking and writing processes needed for independent writing.

2. Literature review

2.1. Definition

2.1.1 Generative AI

The term Generative Artificial Intelligence (GenAI) refers to a branch of artificial intelligence that can produce new content based on patterns learned from large datasets. Unlike traditional AI systems that mainly classify, analyze, or predict existing information, GenAI is characterized by its ability to generate new outputs in different formats, including text, images, audio, video, and computer code, in response to user input (Wang et al., 2024; Xhaferi, 2024). Firstly, GenAI is characterized by its capacity to produce novel content by identifying and drawing on patterns learned from large-scale training datasets (Banh & Strobel, 2023; He et al., 2025). Secondly, as a prompt-based system, GenAI allows users to interact with it through instructions, questions, or requests, from which it generates relevant responses according to the given input (Banh & Strobel, 2023; Chen et al., 2025). Thirdly, text-based GenAI tools are often built on large language models, enabling them to process written language and produce coherent, human-like textual responses (Raza et al., 2025). Finally, because GenAI can create, revise, and transform information, it can support different forms of content production across various contexts (Banh & Strobel, 2023; Storey et al., 2025).

2.1.2. Over-reliance on Generative AI

Over-reliance on Generative Artificial Intelligence (GenAI) refers to an excessive pattern of dependence in which users rely on GenAI tools to a degree that goes beyond normal support or assistance (Passi et al., 2024; Zhai et al., 2024). This form of reliance occurs when users frequently depend on AI-generated outputs to perform tasks, make decisions, or solve problems with limited personal judgment, critical evaluation, or independent effort (Parasuraman & Manzey, 2010; Zhai et al., 2024). It is characterized by the repeated use of GenAI as a primary source of answers or solutions rather than as a supplementary tool, which may lead users to accept generated outputs without sufficient verification or reflection (Passi et al., 2024). Over time, such dependence may reduce users' cognitive engagement, as they become less involved in evaluating information, applying their own knowledge, and making independent decisions (Lee et al., 2025; Zhai et al., 2024). In this sense, over-reliance on GenAI can be defined as a problematic pattern of use in which excessive dependence on AI-generated support may weaken users' independent thinking, judgment, and skill application.

2.1.3. Essay Writing

Essay writing refers to a form of academic writing in which writers develop a central idea, organize supporting points, and present ideas in a clear, coherent, and structured written form (Bailey, 2025; Hamp-Lyons & Heasley, 2006; Hyland, 2019). Unlike informal writing, essay writing follows a structured format that usually includes an introduction, body paragraphs, and a conclusion. It requires writers to state a central idea, develop supporting points, connect ideas coherently, and use appropriate academic language to communicate meaning effectively (Alkamel & Alwagieh, 2024). In second language learning, essay writing is commonly understood as a complex writing task because learners need to combine language proficiency with composing skills during the writing process (Hyland, 2019; Weigle, 2002). This process involves several cognitive activities, including planning, generating ideas, organizing content, translating thoughts into written language, and reviewing or revising written work (Flower & Hayes, 1981; Kellogg, 2008). Therefore, essay writing is not simply the act of producing a written text; rather, it is a recursive process in which learners construct meaning, arrange ideas, and present their thoughts in a coherent academic form (Flower & Hayes, 1981; Hyland, 2019).

2.2. Theorical Framework

2.2.1. Cognitive Load Theory

John Sweller introduced Cognitive Load Theory (CLT) to explain how learners process information during learning. According to Sweller (1988), learners have limited working memory capacity, which means that they can only process a certain amount of information at one time. When a learning task requires too much mental effort, learning may become less effective. Cognitive Load Theory includes three types of cognitive load: intrinsic load, extraneous load, and germane load. Intrinsic load refers to the inherent difficulty of a task, extraneous load results from unnecessary or poorly presented information, and germane load refers to the mental effort devoted to processing, organizing, and constructing knowledge (Sweller, 1988; Sweller et al., 2011).

Essay writing can be viewed as a cognitively demanding task because students need to manage several cognitive and linguistic processes within limited working memory. When writing an essay, they must generate ideas, organize arguments, use appropriate grammar and vocabulary, and maintain coherence across the text (Flower & Hayes, 1981; Jiang & Kalyuga, 2022). These demands may be particularly challenging for EFL students who are still developing both language proficiency and academic writing skills (Nawal, 2018; Jiang &

Kalyuga, 2022). In this context, ChatGPT may reduce part of the immediate cognitive burden by providing language and content support (Patac & Patac, 2025). However, excessive reliance on ChatGPT may become problematic when students depend on the tool to complete demanding stages of essay writing, as this may reduce their engagement in the mental effort needed for idea development, argument construction, and self-revision (Lee et al., 2025; Zhai et al., 2024).

In relation to the present study, Cognitive Load Theory provides a useful framework for explaining why the overuse of ChatGPT may affect students' essay writing development. If students frequently rely on ChatGPT to handle cognitively demanding writing processes, they may reduce the learning-related mental effort needed to develop independent writing skills. Therefore, CLT helps clarify the connection between students' dependence on ChatGPT, their cognitive engagement, and their ability to write essays independently.

2.2.2. Second Language Acquisition in Writing Essay

Second Language Acquisition (SLA) refers to the process through which learners develop the ability to understand and use a language other than their first language. In the context of English essay writing, SLA is closely connected to how learners develop their linguistic knowledge, writing strategies, and ability to express ideas in a second or foreign language. Essay writing is not only a written product but also a learning process in which students apply grammar, vocabulary, sentence structures, and discourse knowledge to communicate meaning in an academic form. Through this process, learners have opportunities to notice language gaps, test their language use, and improve their control of written English (Krashen, 1982; Swain, 1985; Hyland, 2019).

In second language essay writing, students are expected to apply their language knowledge while engaging in key writing processes such as generating ideas, organizing content, developing arguments, and revising their work. To produce an academic essay, they need to use grammar, vocabulary, and sentence structures accurately while also presenting ideas in a coherent and logical way. These activities support second language acquisition because learners actively use the target language to construct meaning rather than only receive language input. From this perspective, essay writing plays an important role in SLA as it allows learners to transform their language knowledge into written communication and gradually improve the accuracy, fluency, and complexity of their English.

SLA in essay writing provides a basis for understanding why students' active involvement in the writing process is important. When students depend too heavily on ChatGPT to generate ideas, organize arguments, or produce written content, they may reduce their opportunities to practice using English independently. This may affect the language-learning value of essay writing, as students become less engaged in applying grammar, vocabulary, argumentation, and revision strategies by themselves. Therefore, this section helps explain why the overuse of ChatGPT in essay writing may influence not only students' writing performance but also their development as second language learners.

2.3. Previous Research

A study conducted by Mahapatra (2024) investigated the impact of ChatGPT on ESL students' academic writing skills through a mixed-methods approach. The findings revealed that ChatGPT helped students improve grammar, vocabulary, and essay organization. In addition, students reported that the tool made idea generation and revision easier during the writing process. However, the study also found that some learners became overly dependent on AI-generated suggestions, which reduced their independent critical thinking and writing

practice. Mahapatra emphasized that although AI tools can support language learning, excessive reliance on them may negatively affect students' ability to develop writing skills independently.

Bašić et al. (2023) explored the use of ChatGPT-3.5 as a writing assistant in students' essay writing tasks. The researchers discovered that ChatGPT could quickly generate coherent and grammatically accurate essays, helping students complete writing tasks more efficiently. Despite these benefits, the study raised concerns about academic integrity, originality, and students' ownership of ideas. The researchers argued that students who relied too heavily on AI-generated content might become less engaged in the writing process and less capable of producing authentic writing independently.

Research by Boudouaia et al. (2024) focused on the effectiveness of ChatGPT-4 in improving EFL students' writing performance. The findings indicated that students viewed ChatGPT-4 as a useful tool for brainstorming ideas, correcting grammar mistakes, and organizing essay structures. Furthermore, the use of AI assistance reduced students' cognitive burden during complex writing tasks. However, the researchers also observed that students who depended excessively on AI support demonstrated lower levels of creativity and active participation in writing activities. The study suggested that AI should be used as a supportive learning tool rather than a substitute for students' own thinking and writing processes.

Yuan and Dao (2026) examined ESL tertiary learners' engagement in ChatGPT-assisted second language writing. According to the findings, ChatGPT improved students' grammar, coherence, and vocabulary use while also reducing anxiety during writing tasks. Nevertheless, the study revealed that some students became passive during the revision process because they relied heavily on AI-generated feedback instead of revising their essays independently. The researchers concluded that self-regulated learning skills play an important role in preventing excessive dependence on AI tools.

In a Vietnamese educational context, Lai and Trinh (2025) investigated the effects of ChatGPT on the academic writing proficiency of English-major undergraduates. The results showed that ChatGPT positively influenced students' grammar accuracy, vocabulary use, and motivation to write in English. Many participants believed that AI tools made writing tasks easier and more convenient. However, the researchers warned that excessive use of AI-generated assistance could weaken students' independent writing abilities and critical thinking skills over time. They emphasized the importance of balancing AI support with independent learning practices.

2.4. Research Gap

Previous studies have shown that generative AI tools such as ChatGPT can positively support students' academic writing by improving grammar, vocabulary, organization, and writing efficiency (Mahapatra, 2024; Lai & Trinh, 2025). In addition, several researchers have found that AI-assisted writing can reduce students' anxiety and cognitive burden during writing tasks (Boudouaia et al., 2024; Yuan & Dao, 2026).

Although many previous studies have discussed the benefits and challenges of using ChatGPT in writing, there are still several gaps in the existing literature. First, most studies mainly focused on the effectiveness of AI tools in improving writing performance rather than specifically examining the issue of over-reliance on generative AI tools. As a result, there is still limited understanding of how excessive dependence on AI may influence students' independent writing abilities, cognitive engagement, and critical thinking skills over time.

Second, many previous studies were conducted in general ESL or EFL contexts outside Vietnam. Only a small number of studies, such as Lai and Trinh (2025), focused on Vietnamese students. However, there is still limited research investigating the impacts of over-reliance on generative AI tools among Vietnamese university students, particularly sophomores majoring in English. Therefore, more research is needed to better understand how students in the Vietnamese educational context use AI tools in essay writing and how this dependence may affect their learning process.

Finally, previous studies often emphasized either the positive or negative effects of ChatGPT separately, while few studies explored both aspects together within the context of academic essay writing. In addition, limited attention has been given to students' personal experiences, perceptions, and behaviors related to AI dependence in writing tasks. Therefore, the current study aims to fill these gaps by exploring the impacts of over-reliance on generative AI tools in essay writing among HUFLIT sophomores, with a focus on students' perceptions, writing experiences, and independent learning abilities.

3. Methodology

3.1. Research design

To collect data from participants in real time without the need for long-term observation, the study was designed as a cross-sectional survey with semi-structured interviews. This non-experimental design's primary advantage is its capacity to record the widespread application of ChatGPT as they naturally arise in the HUFLIT academic setting (Creswell & Creswell, 2018). The researchers were able to observe sophomores' self-reported habits, experiences, and perceptions in their typical learning context by opting not to alter or control the learning conditions (Bryman, 2016).

For the quantitative phase, a structured questionnaire served as the primary instrument for data collection. All items were presented through Likert-scale statements, allowing for a nuanced measurement of students' attitudes toward AI reliance. The questionnaire focused on variables such as dependence on AI-generated content, reduction in independent idea generation, impact on critical thinking and argument development, and academic confidence and writing autonomy. In addition to the questionnaire, qualitative data was collected through semi-structured interview questions. This approach allowed participants to elaborate on their experiences, writing habits, perceived benefits, and worries about using ChatGPT in academic writing, resulting in rich and nuanced insights that quantitative measurements alone could not capture (Creswell and Poth, 2018). The semi-structured method guaranteed uniformity between interviews while also allowing participants to express their thoughts in greater depth and researchers to investigate emerging themes through follow-up questions (Kallio et al., 2016).

The quantitative data was analyzed using SPSS to summarize participants' reactions to ChatGPT reliance in essay writing. Descriptive statistics, such as frequency, percentage, mean, and standard deviation, were employed to depict the general distribution of responses across questionnaire items. Furthermore, Cronbach's Alpha was used to assess the questionnaire scale's internal consistency (Field, 2018). Meanwhile, the qualitative interview data were evaluated thematically to discover common themes and perspectives connected to AI-assisted writing. The study's goal was to provide a clearer and more thorough assessment of the perceived influence of ChatGPT reliance on students' writing progress, critical thinking, and cognitive engagement (Creswell & Plano Clark, 2018).

3.2. Research context

This study was conducted at the University of Foreign Languages - Information Technology (HUFLIT), a higher education institution in Vietnam where academic writing in English plays a central role in the curriculum. Established in 1994, HUFLIT has over three decades of experience in higher education and professional training, providing academic instruction and career-oriented development for numerous generations of students. The university is guided by the educational philosophy of "pursuing happiness and freedom based on morality and wisdom" and aims to develop globally competent citizens equipped with foreign language proficiency, technological literacy, and lifelong learning capabilities. Currently, HUFLIT consists of multiple faculties and offers a wide range of undergraduate and postgraduate programs across various disciplines.

Within this institutional context, the Faculty of Foreign Languages has become one of the university's most prominent academic departments, with more than thirty years of development in language education and training. The faculty offers several language-related programs, including English Language Studies and Chinese Language Studies. The English Language major has attracted many students due to its diverse career opportunities in education, translation and interpretation, communication, business, and international working environments.

In the context of the English Language Studies (ELS) program, academic writing is an important component of the curriculum. Essay writing courses help students develop academic communication, critical thinking, argument development, and professional language use. In addition, many specialized courses require students to complete written assignments that involve independent idea generation, information evaluation, and logical argument construction. Therefore, this context is relevant for investigating how dependence on ChatGPT may influence students' writing development and cognitive engagement.

3.3. Participants

Table 34

Demographic Information of Survey Participants

Variables		Number of participants	Frequency
Major	Interpreting - Translating	18	14,9%
	Business English	31	25,6%
	English Language Teaching	40	33,1%
	Office Administration	11	9,1%
	English for Logistic	10	8,3%
	English - Chinese Bilingualism	11	9,1%
Total		121	100%

The participants in this study consisted of 127 sophomores majoring in English Language Studies in the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). This total includes 121 questionnaire respondents and six additional interviewees who were recruited separately for the qualitative data. The researcher group selected the participants through convenience sampling from the population of English Language majors. The sample included sophomores from six specializations within the English Language program: (1) Interpreting and Translation, (2) Business English, (3) English Language Teaching, (4) Office Administration, (5) English for Logistics, and (6) English - Chinese Bilingualism.

Table 35

Demographic Information of Interview Participants

Variables	Number of interviewees		Frequency
Major	English Language Teaching	3	50%
	Business English	2	33,3%
	English - Chinese Bilingualism	1	16,7%
TOTAL		6	100%
Interview Mode	Face-to-face	4	66,7%
	Online	2	33,3%
TOTAL		6	100%

The research group selected sophomores because they had recently completed essay writing courses, which constitute an important component of the English Language curriculum. At the time of the study, all participants had experience producing different types of academic essays and written assignments required in both writing and specialized subjects. In addition, all participants had prior experience using ChatGPT to support essay writing. During their coursework, they frequently used ChatGPT to generate ideas, create outlines, organize content, improve grammar, and check vocabulary when completing essay assignments assigned by their lecturers. Consequently, all participants are familiar with using ChatGPT as a writing support tool and are considered suitable for providing relevant information about their experiences with ChatGPT-assisted essay writing.

3.4. Research Instrument

Two research tools were used in the study: semi-structured interview questions and a structured questionnaire. These instruments enabled the collection of both quantitative and qualitative data from HUFLIT sophomores. The research group developed instruments based on previous studies on AI-assisted learning, academic writing, critical thinking, and learner autonomy.

The questionnaire consisted of two parts. The first part collected participants' demographic information, including their specialization within the English Language program. The second part contained twelve closed-ended statements that measured students' perceptions of the potential threats of ChatGPT overuse in essay writing. These statements were organized into four areas: dependence on AI-generated content, reduction in independent idea generation, impact on critical thinking and argument development, and academic confidence and writing autonomy. Participants responded to each statement using a five-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree". In addition to the questionnaire, our research group used semi-structured interview questions to obtain more detailed information about participants' experiences and perceptions of using ChatGPT in essay writing. The interview questions focused on sophomores' writing practices, their purposes for using ChatGPT, and their views on the influence of AI-assisted writing on their learning and writing development. The semi-structured format ensured consistency across interviews while allowing participants to elaborate on their experiences and perspectives.

3.5. Data Collection Procedures

For the quantitative phase, the researchers distributed the questionnaire to 128 sophomores majoring in English Language Studies at HUFLIT. Data collection was conducted

over a period of three weeks. At the end of the collection period, 121 valid responses were returned and included in the analysis. Before completing the questionnaire, participants were informed about the purpose of the study and assured that their responses would remain anonymous and confidential. Participation was entirely voluntary.

In the qualitative data collection process, our research group conducted semi-structured interviews with six additional participants who were not part of the questionnaire sample. These participants are also sophomores majoring in English Language Studies at HUFLIT and had prior experience using ChatGPT for essay writing. Among them, three participants specialized in English Language Teaching, two specialized in Business English, and one specialized in English - Chinese Bilingualism. Four interviews were conducted face-to-face, while the remaining two were conducted online via Google Meet.

3.6. Data Analysis Procedures

For the quantitative data, the research group analyzed the questionnaire responses using SPSS. First the data was coded and entered the software. The five-point Likert scale responses were coded from 1 to 5, with 1 representing "Strongly Disagree" and 5 representing "Strongly Agree." The researchers then used descriptive statistics, including frequency, percentage, mean, and standard deviation, to summarize participants' responses regarding the perceived threats of ChatGPT overuse in essay writing. Before further analysis, the reliability of the questionnaire was examined using Cronbach's Alpha. This test was conducted to determine the internal consistency of the questionnaire items.

For the qualitative data, the researchers analyzed the interview responses using thematic analysis. The recorded interviews were first transcribed. The researchers then read the transcripts several times to become familiar with the data. After that, meaningful responses were coded and grouped into recurring themes related to students' experiences, habits, benefits, and concerns when using ChatGPT for essay writing. These themes were used to support and explain the quantitative findings.

Finally, the researchers integrated the quantitative and qualitative results to provide a more comprehensive understanding of the perceived threats of ChatGPT overuse among HUFLIT sophomores. The quantitative findings showed the general tendencies in participants' responses, while the qualitative findings provided deeper explanations of students' personal experiences with AI-assisted essay writing.

4. Findings and Discussion

This part presents and discusses the findings on HUFLIT sophomores' over-reliance on ChatGPT in English essay writing. The four components of the analysis are Reliance on AI-Generated Content, Reduction in Independent Idea Generation, Impact on Critical Thinking and Argument Development and Academic Confidence and Writing Autonomy. Reliability coefficients, mean scores, standard deviations, and frequency distributions are all used in the discussion to find the key patterns in the responses from the sophomores.

Table 36

Reliability Coefficients of the Four Aspects

Aspects	Number of Items	Cronbach's Alpha
Reliance on AI-Generated Content	3	0,52
Reduction in Independent Idea Generation	3	0,57

Impact on Critical Thinking and Argument Development	3	0,46
Academic Confidence and Writing Autonomy	3	0,50

4.1. Reliance on AI-Generated Content

Table 37

Reliance on AI-Generated Content

Statements	Frequency (%)					Mean	Stand Dev
	SD	D	N	A	SA		
1. I often rely on ChatGPT to generate ideas when writing English essays.	5,8	15,7	16,5	24,0	38,0	3,73	1,28
2. I feel it is difficult to write an essay without using ChatGPT.	5,0	13,2	33,9	35,5	12,4	3,37	1,03
3. I tend to use ChatGPT to help plan the structure of my English essays	2,5	7,4	33,1	37,2	19,8	3,64	0,97
Reliance on AI-Generated Content						3,58	

The first aspect examined sophomores' dependence on ChatGPT for generating and organizing essay content. This dimension recorded Cronbach's Alpha value of $\alpha = 0,52$ and an overall mean score of M = 3,58, the highest among the four dimensions. Table 4 shows that students relied on ChatGPT mostly for idea generation, as the statement "I often rely on ChatGPT to generate ideas when writing English essays" received the highest mean score (M = 3,73) and with 62% of respondents choosing "Agree" or "Strongly Agree". Another notable result is that 57% of sophomores agreed or strongly agreed that they used ChatGPT to plan the structures of their essays. These numbers indicate that ChatGPT was mainly used to support the early stages of essay writing, especially idea generation and organization.

The interview data clarify how this dependence appeared in students' actual writing practices. Several participants described ChatGPT as part of the early production stage of essay writing. One student explained that they used ChatGPT "especially when I do not know how to start" because it helped them feel "less stuck" and gave them "a clearer direction." Another participant gave a similar account, stating that "the first thing I do is paste it into ChatGPT to get ideas because starting from a blank page is the hardest part for me." These responses suggest that students did not use ChatGPT only for grammar correction or final editing; rather, it became involved in the generation and organization of essay content from the beginning. This indicates that students' dependence was strongest at the stage where they needed to form initial ideas and structure their essays.

From the perspective of Cognitive Load Theory (Sweller, 1988), students appeared to use ChatGPT primarily to reduce the cognitive demands associated with brainstorming and organizing ideas. While this strategy may lower the mental effort required during the initial stage of writing, frequent dependence on AI-generated ideas may also reduce opportunities for learners to actively construct their own ideas. This finding is generally consistent with previous studies, suggesting that ChatGPT is most beneficial when used as a support tool rather than as the primary source of content generation.

4.2. Reduction in Independent Idea Generation

Table 38

Reduction in Independent Idea Generation

Statements	Frequency (%)					Mean	tand ev
	SD	D	N	A	SA		
4.I feel that frequent use of ChatGPT affects how I generate ideas independently.	5,0	11,7	30,6	28,1	24,0	3,55	,13
5.I usually turn to ChatGPT when I need ideas for an essay	5,8	11,7	34,2	32,5	15,8	3,41	,08
6. I tend to accept ChatGPT's ideas instead of developing my own viewpoints.	6,7	10,0	28,3	30,0	25,0	3,57	,17
Reduction in Independent Idea Generation						3,51	

The reduction in independent idea generation was the subject of the second dimension. This recorded Cronbach's Alpha value of $\alpha = 0,57$ and an overall mean score of M = 3,51, suggesting a moderate tendency toward reduced independent idea development. Table shows that 55% of the respondents agreed or strongly agreed that they tended to accept ChatGPT's ideas instead of developing their own viewpoints (M = 3,57). Similarly, more than half of the sophomores agreed that frequent use of ChatGPT affected how they generated ideas independently (52,1%, M = 3,55). Although the item about turning to ChatGPT when needing essay ideas had a slightly lower mean score (M = 3,41), nearly half of the respondents still agreed or strongly agreed.

The qualitative responses provide a clearer picture of this shift in idea development. Interviewees repeatedly described a movement from self-generated ideas to AI-assisted idea selection. One participant reflected, "Before using ChatGPT, I had to think more carefully about the topic and try to come up with ideas by myself. But now, when I receive an essay topic, I usually ask ChatGPT first.". Another student expressed this more directly "I have been completely outsourcing the brainstorming process.". These comments indicate that ChatGPT may change the starting point of essay writing. Instead of beginning with independent brainstorming, some students first receive AI-generated options and then choose, modify, or assemble them. This helps explain why the survey results show a moderate reduction in independent idea generation because the problem is not that students never revise AI output, but that their initial thinking may increasingly depend on AI suggestions.

Rather than indicating a complete loss of independent thinking, these findings suggest that students increasingly begin the writing process by consulting AI before generating their own ideas. This behavioral pattern may gradually reshape learners' cognitive engagement because AI becomes the starting point of idea generation rather than a supplementary resource. Consequently, the issue appears to concern the sequence of cognitive processing rather than the complete replacement of students' own ideas.

4.3. Impact on Critical Thinking and Argument Development

Table 39

Impact on Critical Thinking and Argument Development

Statements	Frequency (%)					Mean	tand ev
	SD	D	N	A	SA		
7. Overusing ChatGPT weakens my ability to analyze essay topics critically.	3,3	12,4	30,6	28,1	25,6	3,60	,10
8. I find it harder to develop strong arguments without ChatGPT's support.	4,1	13,2	32,2	38,8	11,6	3,40	,00
9. My essays rely more on ChatGPT's reasoning than my own critical thinking.	7,4	8,3	31,4	35,5	17,4	3,47	,10
Impact on Critical Thinking and Argument Development						3,49	

The third dimension explored how students perceived the influence of ChatGPT on their critical thinking and argument development. With Cronbach's Alpha value of $\alpha = 0,46$ and an overall mean score of M = 3,49, the results suggest a moderate level of perceived impact. Among the items in this dimension, the strongest response was found in the statement that overusing ChatGPT weakened students' ability to analyze essay topics critically (M = 3,60) with 53,7% of respondents selecting agree or strongly agree. A similar pattern appeared in the item concerning reliance on ChatGPT's reasoning, where 52,9% of students agreed or strongly agreed that their essays depended more on ChatGPT's logic than on their own critical thinking (M = 3,47).The mean score for the statement about finding it harder to develop strong arguments without ChatGPT also remained above the neutral point (M = 3,40). These results indicate that students recognized ChatGPT's influence not only on essay content but also on the reasoning process behind their writing.

The interview responses provide further insight into how this influence appeared in students' writing practices. Some participants described a decline in their own analytical engagement when ChatGPT was used too heavily. One student admitted, "I feel like it kind of bypasses my critical thinking" explaining that instead of evaluating perspectives independently, they asked ChatGPT to generate "pros and cons" or "counterarguments." Another student similarly stated that their critical thinking was limited to judging "whether the AI's summary is suitable for the essay.". These comments suggest that, for some students, ChatGPT shifted their role from actively constructing arguments to selecting and adjusting AI-generated reasoning. At the same time, the interviews also showed that ChatGPT did not affect all students in the same way. While some used it passively, others used it more critically. One participant explained: "It helps me see the counterarguments so I can make my actual essay way stronger.". This contrast suggests that ChatGPT's impact on critical thinking depends largely on how students engage with the tool. When students rely on it to produce ready-made reasoning, it may reduce independent analysis. However, when they use it to test, question, or refine their arguments, it can become a useful tool for strengthening critical thinking and argument development.

The interview findings also reveal that the impact of ChatGPT is not uniform across learners. Students who actively questioned, evaluated, and modified AI-generated responses

appeared to maintain higher levels of critical engagement than those who accepted the suggestions with minimal reflection. This difference suggests that the educational impact of ChatGPT depends largely on how students regulate their interaction with the technology rather than on the technology itself. From the perspective of Cognitive Load Theory, reducing cognitive effort is beneficial only when learners continue to invest sufficient germane cognitive resources in constructing their own understanding.

4.4. Academic Confidence and Writing Autonomy

Table 40

Academic Confidence and Writing Autonomy

Statements	Frequency (%)					Mean	Stand Dev
	SD	D	N	A	SA		
10. Excessive use of ChatGPT decreases my confidence in my own writing ability.	5,0	9,2	37,8	30,3	17,6	3,46	1 ,05
11. I feel less independent as a writer when I rely too much on ChatGPT.	4,2	15,0	30,0	41,7	9,2	3,37	0 ,99
12. I worry that overusing ChatGPT may negatively affect my long-term writing skills.	5,8	12,4	28,1	32,2	21,5	3.51	1 ,13
Academic Confidence and Writing Autonomy						3,44	

The fourth aspect considered how ChatGPT use relates to students' academic confidence and sense of writing autonomy. This aspect shows Cronbach's Alpha value of α = 0,50 and had the lowest overall mean score among the four criteria (M = 3,44). Although this suggests that the influence of ChatGPT on confidence and autonomy was not as strong as its influence on idea generation or argument development, the results still point to a noticeable concern. More than half of the sophomores (53.7%) agreed or strongly agreed that overusing ChatGPT might negatively affect their long-term writing skills (M = 3,51). In addition, 47,9% of students reported that excessive use of ChatGPT decreased their confidence in their own writing ability (M = 3,46), while 50,9% felt less independent as writers when relying too much on the tool (M = 3,37). These figures suggest that students were aware of the possible impact of ChatGPT on their development as independent academic writers.

The interview data provides a more detailed explanation of this pattern. For many students, ChatGPT was attractive because it made essay writing more efficient and manageable. One participant explained, "I use ChatGPT to maximize my efficiency. With multiple assignments due in the same week, I don't have the luxury to spend days on one paper.". Another student described the writing process with ChatGPT as "much less stressful, much faster.". These responses show why students continue to depend on the tool, because it helped them reduce academic pressure, complete assignments more quickly, and feel more secure about the quality of their work. However, the same convenience also appeared to create a weaker sense of ownership and independence for some students. One participant referred to ChatGPT as a "safety net" suggesting that the tool had become something they depended on for reassurance. A stronger expression of dependence appeared in another response: "Without AI, I have no confidence.". Similarly, one student admitted, "I don't feel like I own the essay.".

These comments reveal that the problem is not only whether ChatGPT helps students produce better essays, but whether students still feel capable of writing without its support. Therefore, while ChatGPT offers clear academic convenience, repeated reliance on it may gradually reduce students' confidence, independence, and sense of authorship in their own writing.

Although ChatGPT helped students complete assignments more efficiently, the qualitative findings indicate that prolonged dependence may gradually weaken learners' confidence in their own writing ability. The issue therefore extends beyond writing performance to students' perceptions of themselves as independent academic writers. This observation is consistent with previous research highlighting that excessive reliance on AI may reduce learners' sense of ownership and autonomy despite improving short-term writing efficiency.

5. Conclusion

5.1. Implications

The aim of the current study is to investigate the perceived threats of overusing ChatGPT in essay writing among HUFLIT sophomores majoring in English Language Studies. Based on the survey and interview results, some implications for teaching and learning are drawn as follows.

Firstly, the strong tendency to rely on ChatGPT in the early stages of essay writing emphasizes the need for students to use the tool as a form of support rather than a replacement for independent thinking. Thus, students should be encouraged to take responsibility for the content, logic, coherence, and personal voice of their essays, even when ChatGPT is used for brainstorming, outlining, grammar checking, or idea refinement. In addition, the habit of critically evaluating and revising AI-generated suggestions may help students avoid accepting ready-made ideas too passively. As a result, students can benefit from AI assistance while still maintaining their independence as academic writers.

In addition, the findings suggest that lecturers play an important role in guiding students toward responsible AI use. Since many students have already included ChatGPT in their writing process, writing instruction should not only focus on the final essay but also on how students generate ideas, evaluate arguments, revise content, and make independent writing decisions. Thus, lecturers need to help students distinguish between using ChatGPT as useful support and depending on it excessively.

5.2. Limitations

Some limitations of the current study should also be acknowledged. First, the study was conducted only with HUFLIT sophomores majoring in English Language Studies, so the findings may not fully represent students from other academic years, majors, or institutions. Second, the questionnaire relied on students' self-reported responses, which means that the results reflected their perceptions rather than direct evidence of their actual writing behavior. Third, the number of interview participants was limited, so the qualitative findings could not cover the full range of students' experiences with ChatGPT in essay writing. In addition, the reliability coefficients of some questionnaire aspects were relatively modest, which suggests that the results should be interpreted as exploratory rather than conclusive.

5.3. Recommendations

Keeping ChatGPT as an enabler rather than a substitute for real-life learning, universities must move from passive uses of it to structured educational frameworks. Firstly,

educators need practical rules to ensure ChatGPT is treated as a basic learning tool, not a shortcut for outsourcing thought. Rather than just grading a completed paper, faculty can structure assignments to require clear evidence of active problem-solving. When students are required to submit independent brainstorming notes, rough drafts, and self-reflection logs, they are forced to stay cognitively engaged. This prevents them from using AI to skip vital critical thinking steps. Furthermore, future studies should look past the cross-sectional framework applied here and transition toward longitudinal methods. Tracking English majors across consecutive semesters would offer far deeper insights into the long-term learning consequences that AI overuse has on writing independence. Following English language students over multiple semesters will offer far deeper insights into the long-term learning consequences that AI dependency has on writing independence and structural language growth.

References

Alkamel, M. A. A., & Alwagieh, N. A. S. (2024). Utilizing an adaptable artificial intelligence writing tool (ChatGPT) to enhance academic writing skills among Yemeni university EFL students. *Social Sciences & Humanities Open, 10*, Article 101095. https://doi.org/10.1016/j.ssaho.2024.101095

Bailey, S. (2025). *Academic writing: A handbook for international students* (6th ed.). Routledge. https://doi.org/10.4324/9781003509264

Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets, 33*, Article 63. https://doi.org/10.1007/s12525-023-00680-1

Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications, 10*, Article 750. https://doi.org/10.1057/s41599-023-02269-7

Boudouaia, A., Mouas, S., & Kouider, B. (2024). A study on ChatGPT-4 as an innovative approach to enhancing English as a foreign language writing learning. *Journal of Educational Computing Research, 62*(6), 1289-1317 https://doi.org/10.1177/07356331241247465

Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.

Chen, B., Zhang, Z., Langrené, N., & Zhu, S. (2025). Unleashing the potential of prompt engineering for large language models. *Patterns (N Y), 6*(6), Article 101260. https://doi.org/10.1016/j.patter.2025.101260

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Creswell, J. W., Clark, V. L. P., & Clark, V. L. (. P. (2017). Designing and conducting mixed methods research. Sage Publications, Inc.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Crosthwaite, P., & Sun, S. (2025). Generative AI and L2 written feedback studies: A scoping review. *RELC Journal*. Advance online publication. https://doi.org/10.1177/00336882251386530

Ding, L., Zou, D., & Kohnke, L. (2026). Enhancing L2 writing with AI: Effects of ChatGPT feedback and the role of student engagement. *Artificial Intelligence in Language Education, 2*, Article 103356. https://doi.org/10.29140/aile.2026.103356

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32(4), 365-365. https://doi.org/10.2307/356600

Hamp-Lyons, L., & Heasley, B. (2006). *Study writing: A course in writing skills for academic purposes* (2nd ed.). Cambridge University Press.

R, H., J, C., & T., T. (2025). Generative artificial intelligence: a historical perspective. Natl Sci Rev, 12(5), nwaf050. https://doi.org/10.1093/nsr/nwaf050

Hyland, K. (2003). *Second language writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667251

Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

Jiang, D., & Kalyuga, S. (2022). Learning English as a foreign language writing skills in collaborative settings: A cognitive load perspective. *Frontiers in Psychology, 13*, Article 932291. https://doi.org/10.3389/fpsyg.2022.932291

Jmaiel, H. A., Abukhait, R. O., Mohamed, A. M., Shaaban, T. S., Al-khresheh, M. H., & Al-Qadri, A. H. (2025). The role of ChatGPT in enhancing EFL students' ESP writing skills: An experimental study of gender and major differences. *Discover Education, 4*, Article 240. https://doi.org/10.1007/s44217-025-00700-6

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi- structured interview guide. *Journal of Advanced Nursing, 72*(12), 2954–2965. https://doi.org/10.1111/jan.13031

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1(1), 1–26. https://doi.org/10.17239/jowr-2008.01.01.1*

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Lai, T. K. L., & Trinh, H. L. (2025). The effects of ChatGPT on academic writing proficiency among English-majored undergraduates at a university. *VNU Journal of Foreign Studies, 41*(3), 92–109. https://doi.org/10.63023/2525-2445/jfs.ulis.5514

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–22). ACM. https://doi.org/10.1145/3706598.3713778

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments, 11*, Article 9. https://doi.org/10.1186/s40561-024-00295-9

Nawal, A. F. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies, 3*(1), 385–402. https://doi.org/10.1080/23752696.2018.1513812

R, P., & DH., M. (2010). Complacency and bias in human use of automation: an attentional integration. Hum Factors, 52(3), 381-410. https://doi.org/10.1177/0018720810376055

Samir Passi, Shipi Dhanorkar, & Mihaela Vorvoreanu. 2024. *Appropriate Reliance on Generative AI: Research Synthesis*. Microsoft Technical Report MSR-TR-2024- 7. Microsoft Corporation.

Patac, L. P., & Patac, A. V., Jr. (2025). Using ChatGPT for academic support: Managing cognitive load and enhancing learning efficiency—A phenomenological approach. *Social Sciences & Humanities Open, 11*, Article 101301. https://doi.org/10.1016/j.ssaho.2025.101301

Raza, M., Jahangir, Z., Riaz, M. B., Saeed, M. J., & Sattar, M. A. (2025). Industrial applications of large language models. *Scientific Reports, 15*, Article 13755. https://doi.org/10.1038/s41598-025-98483-1

Setyowati, L., Sukmawan, S., & El-Sulukkiyah, A. A. (2020). Exploring the use of ESL composition profile for college writing in the indonesian context. International Journal of Language Education, 171-182. https://doi.org/10.26858/ijole.v4i2.13662

Storey, V. C., Yue, W. T., Zhao, J. L., & Lukyanenko, R. (2025). Generative artificial intelligence: Evolving technology, growing societal impact, and opportunities for information systems research. *Information Systems Frontiers, 27*, 2081– 2102. https://doi.org/10.1007/s10796-025-10581-7

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. Springer New York. https://doi.org/10.1007/978-1-4419-8126-4

Wang, C., Li, Z., & Bonk, C. J. (2024). Understanding self-directed learning in AI-assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence, 6*, Article 100247. https://doi.org/10.1016/j.caeai.2024.100247

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732997

Wingate, U. (2012). "Argument!" Helping students understand what essay writing is about. *Journal of English for Academic Purposes, 11*(2), 145–154. https://doi.org/10.1016/j.jeap.2011.11.001

Xhaferi, A. (2024). The integration of ChatGPT in academic writing: Exploring perceptions and effectiveness of the students. *International Journal of Multilingual Education, 25*(2), 74–86. https://doi.org/10.22333/ijme.2024.8997

Yuan, S., & Dao, P. (2026). ESL tertiary learners' engagement in ChatGPT-assisted second language writing. *Language Awareness*. Advance online publication. https://doi.org/10.1080/09658416.2026.2652333

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*, Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education*, *21*(1), 1–14. https://doi.org/10.1186/s41239-024-00467-0

Bionote

Pham Kim Quoc is a junior majoring in English Language Studies, with a specialization in English Language Teaching, at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). His academic interests include English grammar, linguistic analysis, teaching methodology, and academic writing. He is particularly interested in language teaching and the application of generative AI tools in English learning and writing development.

Le Hoang Yen Vy is a junior majoring in English Language Studies, with a specialization in English Language Teaching, at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). Her academic interests include education, linguistic research, academic writing, and technology-assisted language learning. She is particularly interested in how generative AI tools influence students' writing practices and independent learning in higher education.

Tran Nguyen Thu Ha is a junior majoring in English Language Studies, with a specialization in English Language Teaching, at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). Her academic interests include English language teaching, second language learning, lesson planning, and teaching methodology. She is particularly interested in developing effective teaching practices and creating engaging learning activities for English learners.

Bui Hoang Gia Han is a junior majoring in English Language Studies, with a specialization in English Language Teaching, at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT). Her academic interests include English language teaching, learner-centered instruction, lesson planning, and classroom interaction. She is particularly interested in designing communicative learning activities that support student participation and language development.

Ms. Tran Thi Truc Duyen holds an MA in Education (TESOL) and is a lecturer in the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Vietnam. She has more than 15 years of teaching experience in English language education and specializes in English language teaching methodology and Applied Linguistics. Her research interests include English language education, pre-service teacher education, teacher cognition, educational innovation, and lifelong learning within Vietnamese educational contexts.

Appendix A

Questionnaire

Dear Sophomores,

We are currently students of the Faculty of English Language at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT).

At present, we are taking the Research 2 course and have a strong interest in the use of artificial intelligence in education, particularly in the field of essay writing. Therefore, our research team has decided to conduct a study entitled "The Threats of Overusing ChatGPT on HUFLIT Sophomores in Essay Writing."

We would greatly appreciate it if you could spare a few minutes of your valuable time to participate in our research by answering the questions in the survey below.

Your participation in this questionnaire is highly valuable to us. All responses will remain confidential and will only be used for academic purposes We look forward to your valuable cooperation and contributions!

Thank you very much!

I. Personal infomation

1. What is your major?

1. Interpreting - Translating
2. Business English
3. English Language Teaching
4. Office Administration
5. English for Logistic
6. English - Chinese Bilingualism

☐ 1
☐ 2
☐ 3
☐ 4
☐ 5
☐ 6

II. The Threats of Overusing ChatGPT on HUFLIT's Sophomores in Essay Writing

Please indicate your level of agreement with the following statements about overuse of ChatGPT in essay writing

1. Strongly Disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

Section A

Reliance on AI-generated Content

	1	2	3	4	5
1. I often rely on ChatGPT to generate ideas when writing English essays.	☐	☐	☐	☐	☐
2. I feel it is difficult to write an essay without using ChatGPT.	☐	☐	☐	☐	☐
3. I tend to use ChatGPT to help plan the structure of my English essays	☐	☐	☐	☐	☐

Section B:

Reduction in Independent Idea Generation

	1	2	3	4	5
4. I feel that frequent use of ChatGPT affects how I generate ideas independently.	☐	☐	☐	☐	☐
5. I usually turn to ChatGPT when I need ideas for an essay	☐	☐	☐	☐	☐
6. I tend to accept ChatGPT's ideas instead of developing my own viewpoints.	☐	☐	☐	☐	☐

Section C

Impact on Critical Thinking and Argument Development

	1	2	3	4	5
7. Overusing ChatGPT weakens my ability to analyze essay topics critically.	☐	☐	☐	☐	☐
8. I find it harder to develop strong arguments without ChatGPT's support.	☐	☐	☐	☐	☐
9. My essays rely more on ChatGPT's reasoning than my own critical thinking.	☐	☐	☐	☐	☐

Section D

Academic Confidence and Writing Autonomy

	1	2	3	4	5
10. Excessive use of ChatGPT decreases my confidence in my own writing ability.	☐	☐	☐	☐	☐
11. I feel less independent as a writer when I rely too much on ChatGPT.	☐	☐	☐	☐	☐
12. I worry that overusing ChatGPT may negatively affect my long-term writing skills.	☐	☐	☐	☐	☐

Appendix B

Semi-structured Interview Questions

1. How do you usually use Generative AI tools such as ChatGPT when writing essays?

Possible follow-up questions:

- At which stage of essay writing do you usually use ChatGPT? (brainstorming, outlining, writing, editing)

- Do you usually modify ChatGPT's answers or use them directly? Why?

2. Do you think relying too much on ChatGPT affects your ability to develop your own ideas and write independently? Why or why not?

Possible follow-up questions:

- Do you feel less confident writing essays without ChatGPT? Why?

- How has your writing process changed since you started using ChatGPT?

3. How does using ChatGPT influence your critical thinking during the essay-writing process?

Possible follow-up questions:

- Do you usually check whether ChatGPT's information is correct? How?

- Do you compare ChatGPT's answers with other sources such as textbooks or teacher feedback?

4. Do you think using ChatGPT has changed your essay-writing skills? Why or why not?

Possible follow-up questions:

- How different is your writing process now compared to before using ChatGPT?

- Do you still practice writing essays by yourself without AI support? Why or why not?

#159 ENGLISH LISTENING AND NOTE-TAKING: CHALLENGES AND STRATEGIES FOR HUFLIT ENGLISH MAJORS

Bùi Thị Thu Ngân[1]

Abstract

Proficiency in English listening and note-taking skills is fundamental to the English Language Studies curriculum; however, many university students struggle to master this integrated skill. This study investigates the challenges faced by English majors at HUFLIT, examining students' reported coping strategies and the relationship between note-taking difficulties and strategy use. Utilizing a sample of 78 students, the research explores the sequential relationship among listening comprehension barriers, note-taking difficulties, and coping strategies. A 22-item reflective scale comprising these three constructs was developed and validated using SmartPLS. The structural equation model (PLS-SEM) tested two primary hypotheses: the impact of listening challenges on note-taking difficulties (H1) and the influence of those difficulties on strategy adoption (H2). Findings indicate that listening barriers—such as unfamiliar vocabulary, delivery speed, and lack of background knowledge—significantly predict note-taking struggles; however, these struggles do not have a significant impact on the implementation of coping mechanisms. While students report using various strategies, including pre-class preparation, symbols, abbreviations, and the integration of AI tools, the adoption of these strategies is not directly driven by note-taking difficulty. Practical recommendations are provided for students, instructors, and curriculum developers to overcome challenges in English Listening and Note-Taking classes.

***Keywords*:** Listening challenges, note-taking, English majors, coping strategies

1. Introduction

Listening and note-taking are essential academic skills that enable students to understand, record, and review information from lectures. For English majors, these skills are particularly important because they support learning in content-based courses and later professional subjects, especially English for Specific Purposes (ESP) and Interpretation courses.

At Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Listening and Note-Taking is a compulsory course for English-majored freshmen. Before taking this course, students complete a Listening and Speaking course to develop their basic language skills. The Listening and Note-Taking course then provides explicit training in listening comprehension and note-taking techniques through a blended learning approach that combines face-to-face and online instruction.

Despite recognizing the importance of these skills, many freshmen continue to experience difficulties when listening to lectures and taking notes in English. Common challenges include understanding specialized vocabulary, following fast speech, maintaining concentration, identifying key information, and recording notes efficiently. These difficulties

[1] *HUFLIT - HCMC University of Foreign Languages and Information Technology*
Email: nganbtt@huflit.edu.vn
Phone number: 0903509160

may negatively affect students' academic performance and their readiness for more demanding courses in later years.

Previous studies have highlighted the importance of listening comprehension and note-taking in academic success. However, most studies have examined listening difficulties, note-taking techniques, or learning strategies separately. Limited research has investigated how listening challenges lead to note-taking difficulties and how students cope with these problems within a single framework. In addition, evidence from Vietnamese English-majored freshmen remains scarce particularly in the context of Listening and Note-Taking courses.

To address these gaps, the present study adds value by moving beyond a purely descriptive analysis to examine whether listening barriers directly lead to note-taking problems, and subsequently, whether these problems actually trigger coping strategies among English-majored freshmen at HUFLIT. By testing this sequential relationship in one comprehensive model, this study establishes a rationale for understanding students' learning barriers. Ultimately, these findings are expected to provide actionable insights for students, instructors, and curriculum developers to improve integrated listening and note-taking instruction.

2. Literature review

Cognitive pressure from listening input and lecture complexity

Listening skill is one of the most vital but demanding language skills for English Linguistic Studies (ELS) learners because it requires rapid perception, interpretation, and integration of spoken input in real time. Learners must process vocabulary, pronunciation, syntax, discourse markers, and speaker intention simultaneously, which places a heavy cognitive burden on them. Prior research shows that common listening problems include fast speech rate, unfamiliar accents, difficult vocabulary, long utterances, implied meanings, and unfamiliar topics. In the Vietnamese context, these problems are often intensified by limited exposure to authentic spoken English and reduced opportunities for interactive listening practice (Tran & Duong, 2020).

This cognitive load is further worsened by lecture complexity. Lectures that contain dense information, abstract concepts, technical terminology, or an unclear structure dramatically increase cognitive pressure. Students may understand individual words but still fail to grasp the overall message when the lecture is too demanding. For English-majored students, this issue is especially pronounced because classroom lectures often contain both language learning and subject-specific knowledge (Phuong, Lam & Le, 2023). Consequently, the primary challenge is not solely decoding the language, but managing the overwhelming cognitive pressure generated by the listening input.

Effect of cognitive pressure on note-taking performance

The cognitive pressure generated by complex listening input directly impacts a student's ability to execute secondary academic tasks, most notably note-taking. Note-taking is an indispensable academic skill and a cognitive tool that supports comprehension, memory, post-listening review, and interpretation processes. Effective note-taking requires students to identify main ideas, organize key points, and use abbreviations or symbols efficiently. Research on note-taking techniques among English majors in Viet Nam indicates that content organization, visual aids, symbols, language use, abbreviations, and note-taking speed are all components of successful performance, with content organization having a particularly strong influence on listening comprehension (Vo, 2026).

However, in Listening and Note-Taking classes, the demand of processing spoken input and generating written output becomes highly problematic. English-majored students must listen carefully while simultaneously selecting, organizing, and recording information. When the cognitive pressure from lecture speed or unfamiliar vocabulary exceeds the student's processing capacity, they lose track of the speaker's message and subsequently fail to take useful notes. Likewise, when lecture content is dense or unclear, students struggle to identify what to write down. Therefore, listening barriers and note-taking problems are not isolated phenomena; rather, listening difficulties act as a primary barrier that sequentially triggers note-taking failure (Vo, 2026).

Self-regulated coping strategies as a behavioral response

Faced with these listening and note-taking difficulties, students must respond behaviorally to manage their learning. To mitigate these dual challenges, students often adopt - or fail to adopt - a variety of self-regulated coping strategies. These strategies operate as a direct response to the cognitive overload they experience in class. Preventative and active strategies include previewing lesson materials before class, using a shorthand system (e.g., abbreviations and symbols) during lectures, focusing on keywords instead of full sentences, reviewing notes after class, and comparing notes with classmates (Suritsky, 1993).

Furthermore, learners may employ memory-based strategies such as repetition, chunking, mnemonics, and spaced review to strengthen recall of lecture content. These behaviors reflect self-regulated learning, as students actively plan, monitor, and adjust their learning process to overcome note-taking breakdowns. Collaborative and structured interventions also facilitate this response; for instance, studies on Dictogloss - a grammar dictation technique developed by Wajnryb (1990, as cited in Tran, n.d.) where students listen, take notes, and collaborate to reconstruct the text - demonstrate that structured practice can help students develop these vital coping mechanisms to improve both listening and note-taking performance (Nguyen & Nguyen, 2023).

In sum, the literature indicates that listening challenges, lecture complexity, note-taking difficulties, and coping strategies are deeply interconnected. Rather than operating independently, the literature suggests a sequential framework: listening input creates intense cognitive pressure, this pressure subsequently disrupts note-taking performance, and the resulting difficulties ultimately determine whether students adopt self-regulated coping strategies.

Research gaps

Although previous studies have demonstrated the importance of listening comprehension and note-taking skills in academic learning, most have examined these factors separately. Existing research has largely focused on either listening difficulties or effective note-taking techniques, with limited attention to how listening challenges contribute to note-taking difficulties and how students respond through coping strategies. As a result, the relationships among these variables remain insufficiently understood.

Furthermore, previous studies have mainly emphasized effective note-taking practices and their impact on listening performance. Less attention has been paid to the specific listening barriers that hinder note-taking, such as fast lecture speed, unfamiliar vocabulary, lack of background knowledge, concentration loss, and listening fatigue. Similarly, note-taking difficulties, including slow writing speed, difficulty identifying main ideas, information overload, and mental translation, have received relatively limited investigation.

Another important gap concerns students' coping strategies. While some studies have explored traditional learning strategies, little research has examined how students cope with listening and note-taking difficulties in English-medium learning environments. In particular, the role of emerging technology-based strategies, such as AI-assisted note organization, remains under-explored.

In the Vietnamese context, empirical evidence is still limited, especially among freshman English majors who are transitioning from secondary school learning to university-level academic listening. Few studies have investigated how listening challenges, note-taking difficulties, and coping strategies interact within a single framework. Moreover, most previous studies have relied on descriptive or regression-based analyses, whereas research using Structural Equation Modeling (SEM) to examine the relationships among these constructs remains scarce.

To address these gaps, the present study investigates the relationships among Listening Challenges, Note-Taking Challenges, and Coping Strategies among freshman English majors in Viet Nam using a PLS-SEM approach. By integrating these variables into a single model, the study aims to provide a more comprehensive understanding of students' difficulties and adaptive behaviors in Listening and Note-Taking courses.

Conceptual framework

This study proposes that English-majored university students' listening challenges affect their note-taking difficulties, and these note-taking difficulties, in turn, influence the coping strategies they use during and after Listening and Note-Taking classes. In an academic Listening and Note-Taking class, students are required to process long lectures on a variety of topics in English, identify key information, and simultaneously record it efficiently in an outline format. This dual processing makes note-taking an essential but highly demanding skill. When listening input becomes difficult due to speed, accent, advanced vocabulary, unfamiliar topics, or high lecture complexity, students naturally struggle to keep up with the input and produce effective notes (Borkovska & Irgin, 2026).

The model assumes that students do not remain passive when facing these cascading challenges. Instead, they attempt to adopt strategies - such as previewing materials, using abbreviations, reviewing notes, and employing memorization techniques - to cope with the class demands. However, while the model proposes a sequential structure (Listening Challenges $\rightarrow$ Note-Taking Difficulties $\rightarrow$ Coping Strategies), the framework acknowledges that the adoption of coping strategies is not strictly linear. Drawing on self-regulated learning theory, a student's behavioral response to academic pressure does not depend solely on the severity of their note-taking difficulties. The actual deployment of coping strategies is multifaceted and influenced by external and internal resources. Factors such as intrinsic motivation, digital literacy, previous strategy training, and the level of explicit teacher guidance play a critical role in determining whether a student successfully adopts a strategy when faced with a challenge (Permana & Cunandar, 2025).

Hypotheses development

Listening challenges and note-taking difficulties

Listening comprehension is a prerequisite for effective note-taking; students must first decode and understand spoken input before they can accurately record it. If learners cannot catch important words, phrases, or ideas due to fast speech, unfamiliar accents, or limited vocabulary, their note-taking quality is highly likely to decline. Previous research indicates that

listening problems and note-taking problems are closely intertwined in lecture-based learning, particularly in contexts where learners have limited exposure to authentic spoken English. Based on this reasoning, the first hypothesis is proposed:

H1: Listening challenges have a positive effect on note-taking difficulties.

Note-taking difficulties and coping strategies

When students encounter friction in note-taking, they generally attempt to respond by adopting strategies that help manage or reduce these academic burdens. These may include using symbols, abbreviations, structured formats, post-class review, peer comparison, or memorization techniques. Research on meta-cognitive listening strategies indicates that strategy use logically becomes more necessary when learners face dense or demanding lecture content. However, because strategy adoption is also highly dependent on a student's self-regulated learning capacity, motivation, and prior training, the transition from experiencing a difficulty to successfully executing a strategy is complex. Despite the presence of these external mitigating factors, the baseline theoretical expectation is that increased academic pressure acts as the primary trigger for a strategic response. Therefore, the second hypothesis is:

H2: Note-taking difficulties have a positive effect on coping strategies.

In summary, the proposed framework suggests that listening challenges create difficulties in note-taking, which in turn act as the catalyst for coping strategies. The model is highly suitable for PLS-SEM because it allows for the simultaneous examination of multiple relationships among latent constructs, testing both direct and indirect effects. This structure reflects the academic reality of English-majored freshmen in Viet Nam, who must manage spoken input and complex content while navigating their own varying levels of self-regulation and learning autonomy.

3. Methodology

This study employed a quantitative, cross-sectional research design to examine the relationships among listening challenges, note-taking difficulties, and coping strategies among English-majored university students at HUFLIT, Viet Nam. A quantitative approach was selected because it enables the measurement of latent constructs and the statistical testing of structural hypotheses. Data were collected at a single point in time from students who had recently completed a compulsory Listening and Note-Taking course (Chris, 2021).

Research design

The study utilized an explanatory and predictive design, mapping the sequential impact of academic challenges on student behavior. The conceptual model was analyzed using PLS-SEM. PLS-SEM was chosen over simple multiple regression because regression cannot simultaneously handle latent variables with multiple indicators or account for measurement error. Furthermore, PLS-SEM evaluates both the measurement model (the reliability and validity of the indicators) and the structural model (the relationships between constructs) concurrently. This variance-based approach is particularly well-suited for exploratory, prediction-oriented research and remains statistically robust even with small-to-moderate sample sizes and non-normally distributed data (Hair et al., 2011; Jansen, 2025).

Participants and sampling

The target population consisted of English majors at HUFLIT. Before taking the course, these students complete a prerequisite Listening and Speaking class designed to build foundational skills at the B1-B2 proficiency level. The Listening and Note-Taking course then

provides explicit training in essential academic skills necessary for advanced, content-heavy English for Specific Purposes (ESP) and Interpretation courses. The course utilizes a blended learning approach, combining on-campus formal instruction with online practice.

A convenience sampling method was used, drawing from two class groups. The survey achieved a 100% response rate, yielding 78 valid responses. The sample was predominantly female (55 female, 23 male). Most participants were freshmen (92.3%), alongside a small percentage of juniors (3.8%) and seniors (3.8%) fulfilling the requirement. All 78 participants had completed the exact same Listening and Note-Taking course.

While convenience sampling is highly practical for accessing a specific cohort, it presents a limitation in representativeness; findings from this specific HUFLIT cohort may not generalize to all English majors across Viet Nam. Regarding sample size, while N = 78 is relatively small, it satisfies the minimum threshold for PLS-SEM estimation for a simple model. According to the "5-times rule," the minimum sample size should be 5 times the items of the scale. With only two main paths and a maximum of one predictor pointing to any endogenous construct, 78 responses provide sufficient statistical power for this preliminary exploratory analysis.

Research instrument

Data were collected using a structured, bilingual (English-Vietnamese) questionnaire to ensure clarity and minimize misinterpretation. The initial instrument consisted of a 22-item reflective scale adapted primarily from Crawford's (2015) research on note-taking skills in EFL contexts, supplemented by established theoretical frameworks to ground each construct:

- Listening Challenges: Items were developed based on cognitive load theory and established listening comprehension barriers (e.g., fast speech rate, unfamiliar vocabulary, and accent variations).

- Note-Taking Challenges: Items were drawn from academic listening research, focusing on the dual-task friction of processing input while organizing output (e.g., writing speed limitations, missing key ideas, and the burden of mental translation).

- Coping Strategies: Items were initially rooted in self-regulated learning and metacognitive strategy theory, encompassing both traditional methods (abbreviations, symbols) and active review behaviors.

Responses were recorded using a 5-point Likert scale ranging from 1 (Never) to 5 (Always), which is optimal for measuring the frequency of challenges encountered and strategies utilized.

Data analysis and instrument refinement

The collected data were analyzed using SmartPLS 3. The analysis followed the standard two-stage PLS-SEM procedure.

First, the measurement model was assessed for reliability and validity. Construct reliability was evaluated using Cronbach’s alpha and Composite Reliability (CR). Convergent validity was assessed via the Average Variance Extracted (AVE) and outer item loadings, while discriminant validity was confirmed using the Heterotrait-Monotrait (HTMT) ratio. Variance Inflation Factors (VIF) were checked to rule out collinearity issues.

Construct Refinement: During the measurement model assessment, several initial items measuring broad coping strategies demonstrated poor outer loadings and had to be removed to

achieve acceptable construct validity (AVE > 0.5). Consequently, the finalized Coping Strategies construct relies predominantly on items related to digital and post-listening review - specifically, the use of AI-assisted note organization and comparing notes with transcriptions. While this refinement narrows the construct from the broader range of traditional strategies discussed in the wider literature, it accurately captures the specific, technology-assisted, self-regulated learning behaviors most actively utilized by this modern cohort to mitigate cognitive overload.

In the second stage, the structural model was evaluated to test the hypothesized relationships. Bootstrapping was applied to determine the statistical significance (p-values and t-statistics) of the path coefficients.

Ethical considerations

Ethical principles were strictly maintained throughout the data collection process. Prior to the survey, participants were fully informed regarding the study's academic purpose, the voluntary nature of their involvement, and their right to withdraw at any time without penalty. To encourage honest reporting and mitigate social desirability bias, the survey was completely anonymous; no personally identifiable information (such as names, student IDs, or email addresses) was recorded. All data were aggregated to ensure participant privacy.

Findings

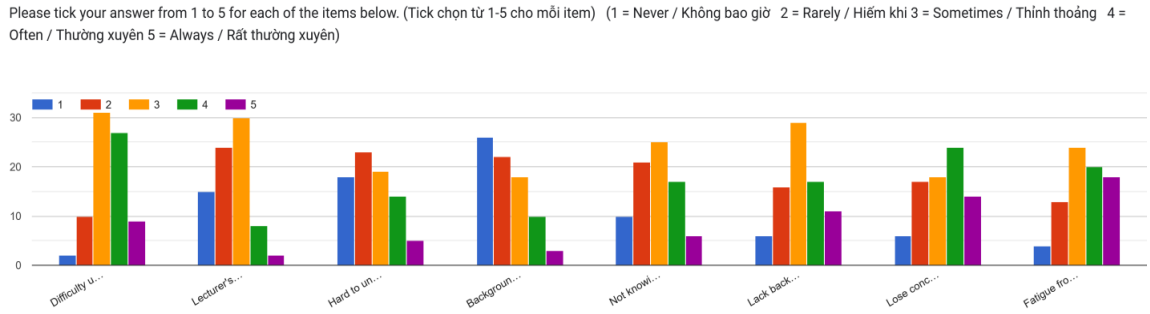
This study examined the relationships among Listening Challenges (LC), Note-Taking Challenges (NTC), and Coping Strategies (CS) among students learning through English-medium lectures.

4. Descriptive Results

The frequency distributions from the survey responses reveal clear patterns regarding the academic hurdles students face during English-medium lectures.

Figure 1

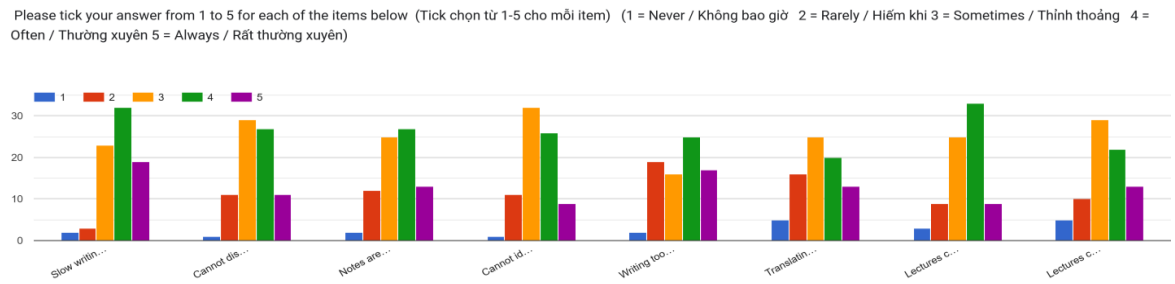
Frequency of experiencing listening challenges among freshman English majors



Listening Challenges: Students frequently experienced barriers related to cognitive overload and linguistic complexity. The highest reported frequencies (clustering heavily in the "Sometimes" and "Often" categories) were attributed to the lecturer's fast speed, specialized vocabulary, lack of background knowledge, concentration loss, and listening fatigue. Conversely, environmental factors such as background noise were reported as less serious problems, with many students indicating they "Never" or "Rarely" struggled with it.

Figure 2

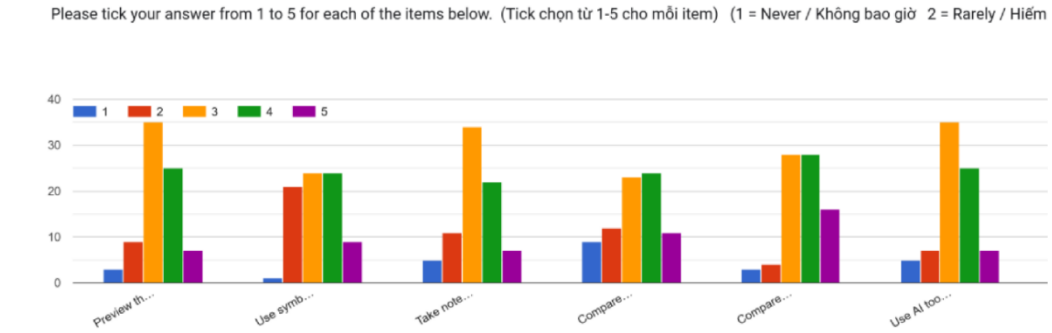
Frequency of experiencing note-taking challenges among freshman English majors



Note-Taking Challenges: The dual-task nature of lectures heavily impacted students. A majority of respondents indicated they "Often" or "Sometimes" struggled with slow writing speed, identifying main ideas, and information overload. Additionally, the cognitive burden of mental translation and general note disorganization were highly prevalent issues.

Figure 3

Frequency of applying coping strategies among freshman English majors



Coping Strategies: While students reported using various strategies, visual data indicates high frequencies for specific technology-assisted methods, particularly the use of AI tools for note organization. However, as detailed in the measurement model results below, the nature of these reported strategies required conceptual refinement.

Measurement Model Results

Construct Reliability and Validity

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
CS		1.000		
LC	0.866	0.884	0.894	0.518
NTC	0.923	0.927	0.937	0.652

Discriminant Validity

Fornell-Larcker Criterion

	CS	LC	NTC
CS			
LC	0.206	0.720	
NTC	0.225	0.790	0.808

Heterotrait-Monotrait Ratio (HTMT)

	LC	NTC
LC		
NTC	0.854	

Collinearity Statistics (VIF)

Outer VIF Values

	VIF
CS5	1.093
CS6	1.093
LC1	1.680
LC2	2.270
LC3	2.228
LC4	1.539
LC5	1.827
LC6	2.447
LC7	2.428
LC8	2.683
NTC1	2.864
NTC2	2.657
NTC3	2.461
NTC4	3.441
NTC5	3.103
NTC6	2.106
NTC7	2.979
NTC8	2.541

The measurement model was assessed to ensure the reliability and validity of the constructs. The reliability and validity results demonstrated that the LC and NTC constructs met and exceeded acceptable standards. The LC construct achieved a Composite Reliability (CR) of 0.894 and an Average Variance Extracted (AVE) of 0.518. Similarly, the NTC construct demonstrated excellent internal consistency and convergent validity with a CR of 0.937 and an AVE of 0.652. Discriminant validity was established, as the Heterotrait-Monotrait (HTMT) ratio between LC and NTC was 0.854, remaining below the conservative 0.90 threshold. All Variance Inflation Factor (VIF) values were below 5, indicating no serious multicollinearity within the model.

Conceptual Refinement of the Coping Strategies (CS) Construct:

During the measurement model assessment, four traditional strategy indicators (CS1, CS2, CS3, and CS4) failed to meet the required outer loading thresholds and were subsequently removed to achieve construct validity. Only two indicators were retained:

CS5: Comparing notes with transcriptions to identify missing points

CS6: Using AI tools to organize notes

Of these, using AI tools was the dominant indicator (loading = 0.935, $p < 0.001$), whereas comparing notes with transcriptions was weaker and non-significant (loading = 0.612, $p = 0.093$). Crucially, the removal of the four traditional items alters the theoretical meaning of this construct. Rather than representing "general coping strategies" as originally hypothesized, the finalized CS construct specifically reflects technology-supported and transcription-based strategies.

Structural Model Results

Path Coefficients

Mean, STDEV, T-Values, P-Values

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values
LC -> NTC	0.790	0.796	0.039	20.239	0.000
NTC -> CS	0.225	0.250	0.134	1.683	0.093

Quality Criteria

R Square

	R Square	R Square Adjusted
CS	0.051	0.038
NTC	0.624	0.619

f Square

	CS	LC	NTC
CS			
LC			1.657
NTC	0.053		

The structural model was evaluated to test the two primary hypotheses.

Hypothesis 1: The Impact of Listening Challenges on Note-Taking

The structural analysis provided robust support for Hypothesis 1. Listening Challenges exhibited a strong, positive, and highly significant effect on Note-Taking Challenges ($\beta = 0.790$, $t = 20.239$, $p < 0.001$). The R-square value for NTC was 0.624, indicating that listening barriers alone account for 62.4% of the variance in a student's note-taking difficulties, representing substantial explanatory power. The effect size was correspondingly very large (f-square = 1.657), confirming that listening comprehension is a critical predictor of note-taking failure.

Hypothesis 2: The Impact of Difficulties on Strategy Adoption

Conversely, Hypothesis 2 was not supported. The relationship between Note-Taking Challenges and technology-supported Coping Strategies was statistically non-significant ($\beta = 0.225$, $t = 1.683$, $p = 0.093$). Furthermore, NTC explained only 5.1% of the variance in CS (R-square = 0.051).

Rather than viewing this solely as a rejected hypothesis, this tension represents a vital finding: the severity of a student's note-taking difficulty does not explain or predict their use of technology-supported coping strategies. Simply experiencing academic hardship is insufficient to trigger strategic adaptation. This suggests that the adoption of AI and transcription tools is likely driven by external factors not captured in this model - such as a student's digital literacy, intrinsic motivation, learning habits, or prior technological training - rather than acting as a direct, reactive response to academic pressure.

5. Discussion

Listening Challenges

The descriptive results showed that students faced several listening difficulties during English-medium lectures. The most frequently reported challenges were fast lecture speed, specialized vocabulary, lack of background knowledge, concentration loss, and listening fatigue. These findings suggest that many students struggled to process spoken information efficiently while maintaining attention throughout the lecture. Among these challenges, fast speech and unfamiliar vocabulary may place a heavy burden on students' working memory. Instead of focusing on the overall message, students may spend too much time decoding

individual words and phrases. Similarly, limited background knowledge makes it more difficult for students to connect new information with what they already know, reducing comprehension. Difficulties with different English accents further increase the complexity of lecture processing. In contrast, background noise appeared to have a relatively minor influence, indicating that linguistic and cognitive factors were more problematic than environmental factors.

These descriptive insights align tightly with the structural equation modeling (SEM) findings, which demonstrate that Listening Challenges strongly and significantly predict Note-Taking Challenges ($\beta = 0.790$, $p < .001$). Because the present study utilizes a cross-sectional design with self-reported data, this relationship suggests a highly probable pathway rather than absolute causality: when students cannot fully decode spoken lecture content, they are subsequently less able to identify key information and record it accurately.

Note-Taking Challenges

The results revealed that students experienced substantial note-taking difficulties. Slow writing speed was one of the most common problems, suggesting a strong association between the pace of the lecture and the physical or cognitive inability to keep up. Students also reported difficulty identifying main ideas and distinguishing them from supporting details, which likely leads to incomplete or ineffective notes. Another important issue was poor note organization. Many students did not use structured note-taking methods such as outlines, Cornell notes, or mind maps. As a result, their notes may have lacked clarity, making later review more difficult. In addition, many participants reported translating English into Vietnamese while listening. This additional mental process is associated with reduced listening comprehension and note-taking. Ultimately, the high explanatory power of the structural model (R-square = 0.624) confirms that listening barriers are the primary statistical driver of note-taking problems.

Coping Strategies

During the analysis, the construct was reduced to only two items - comparing notes with transcriptions and using AI tools to organize notes - with the transcription item yielding a non-significant loading. Consequently, the findings discussed here do not reflect a broad spectrum of traditional coping strategies, but rather a narrow reliance on specific, technology-supported interventions, primarily AI-assisted note organization. This reflects the growing, yet specific, role of educational technology in supporting academic learning. Crucially, the SEM findings indicate that Note-Taking Challenges do not statistically predict the use of these technology-supported Coping Strategies ($\beta = 0.225$, $p = .093$). This indicates that the adoption of coping behaviors is largely independent of the difficulty level experienced by the student. Simply encountering academic obstacles does not automatically trigger strategic adaptation. Instead, strategy use likely depends on independent external or internal factors, such as self-regulation, motivation, digital literacy, previous strategic training, or awareness of available AI tools.

Overall, the findings demonstrate that listening challenges are the main driver influencing note-taking performance among English-majored students. Difficulties such as fast lecture speed and unfamiliar vocabulary create cognitive pressure that is strongly associated with subsequent note-taking difficulties. Simultaneously, the structural results indicate that note-taking challenges alone do not push students to adopt coping strategies. Because strategy adoption - particularly regarding digital and AI-supported tools - appears disconnected from the severity of the student's academic struggle, educators cannot assume students will develop these strategies independently. This highlights a critical need for explicit training. From a

practical perspective, interventions should serve a dual purpose: first, reducing initial listening barriers (e.g., providing pre-lecture vocabulary support and activating background knowledge) before they develop into note-taking problems; and second, explicitly teaching students how and when to apply effective, technology-supported learning strategies to manage English-medium lectures.

Recommendations

The structural and descriptive findings of this study offer clear pathways for pedagogical intervention. Because listening challenges are the primary statistical driver of note-taking difficulties, and because students do not automatically adopt coping strategies simply out of academic need, a multifaceted approach is required.

For Freshman English Majors

a. ***Shift from transcription to keyword systems:*** The descriptive data highlighted slow writing speed and the cognitive burden of mental translation as primary hurdles. To reduce working memory overload, students must abandon full-sentence transcription. Instead, they should adopt a personalized system of abbreviations, symbols, and keywords, which allows them to capture main ideas without falling behind the lecturer's pace.

b. ***Build academic listening stamina:*** To counter the highly reported issues of concentration loss and listening fatigue, students should build their cognitive endurance progressively. Rather than consuming lengthy lectures immediately, students should practice with authentic materials (e.g., TED Talks, university lectures) in short, 10-to-15-minute intervals. This progressive exposure to diverse accents and authentic speeds will better prepare them for the demands of future ESP courses.

c. ***Engage in strategic, verified AI use:*** While AI-assisted note organization was the most prominent coping strategy, passive reliance on these tools undermines long-term retention. Students must use AI ethically and critically. They should treat AI-generated summaries not as final products, but as comparative tools—cross-referencing AI outputs against their own handwritten notes and available transcripts to actively identify knowledge gaps and rectify misunderstandings.

For Instructors

Anticipate cognitive overload: The strongest finding of this study (H1) confirms that listening barriers directly cause note-taking failure. Instructors cannot expect high-quality notes if the foundational listening input is too dense. Lecturers should proactively reduce cognitive pressure by pre-teaching specialized vocabulary, activating background knowledge, and explicitly highlighting signposting language before the lecture begins.

Scaffold the demand: Because students reported severe information overload, instructors should avoid continuous, unbroken presentations. Chunking lectures into 20-minute segments interspersed with brief comprehension checks, discussions, or pauses will give students the necessary time to process information and reorganize their notes without missing new input. Providing partially completed outlines before class can also scaffold this dual-task demand.

Provide explicit strategy and technology training: The non-significant finding for H2 indicates that students do not naturally adopt coping strategies just because they are struggling. Therefore, instructors must explicitly model structured note-taking methods (e.g., Cornell notes, mind mapping) in class. Furthermore, instructors should actively demonstrate how to

use AI tools for post-lecture review, teaching students the difference between using tech for academic support versus academic short-cutting.

For Curriculum Developers

Establish academic support: Because freshmen are navigating a difficult transition to English-medium lectures and do not automatically develop self-regulated learning strategies, institutions should provide more support for them. Implementing peer tutoring, listening workshops, or note-taking clinics can provide the explicit, step-by-step guidance students need outside of regular class hours.

Formalize digital and AI literacy: The findings show that students are already heavily relying on AI to cope with academic demands. Rather than ignoring or simply banning these tools, curriculum developers must formally integrate AI and digital literacy into foundational courses. Teaching students to evaluate, verify, and ethically utilize educational technology ensures that these tools enhance, rather than replace, essential cognitive skills.

Limitations

This study has several limitations that should be considered when interpreting the findings.

Sampling limitations: The study relied on a relatively small sample size (N = 78) drawn from a single institution (HUFLIT) in Viet Nam. While this sample size meets the minimum threshold for preliminary PLS-SEM analysis, it limits the overall statistical robustness of the model and restricts the generalizability of the findings to English majors at other universities or across different proficiency levels.

Measurement limitations: The Coping Strategies construct underwent significant refinement during the measurement model assessment. Because several items were removed to achieve validity, the finalized construct relies heavily on a narrow set of technology-supported and transcription-based indicators. Consequently, the model may not fully capture the complete, broader range of traditional coping strategies that students utilize in practice.

Design and data collection constraints: The cross-sectional design captures data at only a single point in time, which precludes the establishment of absolute causal relationships. Furthermore, the reliance on self-reported questionnaire data introduces the potential for response bias, as students' perceived difficulties may differ from their actual cognitive processing and academic execution.

Omitted variables: The conceptual model did not account for other potentially relevant independent variables, such as student motivation, language anxiety, prior strategy training, or baseline digital literacy. These external factors likely play a significant role in influencing both academic performance and the adoption of coping strategies.

Recommendations for future studies: To address these methodological and conceptual limitations, future research should aim to replicate this structural model using a larger and more geographically diverse sample. Researchers should also move beyond self-reported surveys by incorporating objective listening and note-taking performance tasks to measure actual academic competency. Finally, experimental or longitudinal studies should be conducted to test whether explicit strategy instruction objectively improves students’ long-term listening and note-taking performance.

Implications

Theoretical implications

This study contributes to the literature by demonstrating that academic listening and note-taking must be understood as an integrated cognitive process. Specifically, the findings establish a strong sequential link between listening challenges and note-taking difficulties. However, by revealing that the path from note-taking difficulties to coping strategies is not statistically significant, the study challenges the assumption that these adaptive behaviors are an automatic, sequential response to academic hardship. Instead, it refines current theoretical models by highlighting that while listening input barriers directly disrupt note-taking output, a student's decision to deploy coping strategies operates independently of the academic pressure they experience.

Practical implications

Practically, these findings indicate that course designs cannot rely on students naturally developing adaptive behaviors simply because they are struggling with the material. Therefore, effective pedagogy in English-medium environments requires a dual approach. First, instructors must provide proactive listening support (e.g., pre-teaching vocabulary, activating background knowledge) to reduce the initial cognitive load before it grows into note-taking failure. Second, institutions must implement explicit strategy training - particularly regarding the critical and effective use of digital and AI tools - ensuring students are actively taught how to manage complex classroom input rather than assuming they will adapt independently.

References

Borkovska, N. & Irgin, P. (2026). The effect of structured and conventional note-taking on lecture listening comprehension and specialized vocabulary recall in EMI academic contexts, English for Specific Purposes, Volume 83, 2026, Pages 115-128, ISSN 0889-4906,

https://www.sciencedirect.com/science/article/pii/S0889490626000153

Chris, M. J. (2021). Methodology section for research papers. Retrieved from:

https://www.sjsu.edu/writingcenter/docs/handouts/Methodology.pdf

Crawford, M. J. (2015). A study on note taking in EFL listening instruction. In P. Clements, A. Krause, & H. Brown (Eds.), JALT2014 Conference Proceedings. Tokyo: JALT.

Jansen, D. (2025). How to write the methodology chapter. Retrieved from: https://gradcoach.com/how-to-write-the-methodology-chapter/

Nguyen, L. B. N., Nguyen, H. P. M. (2023). Vietnamese EFL Students' Perception on Using Dictogloss in Listening and Note-taking Classes. Proceedings of the AsiaCALL International Conference, 4, 35-

44, ISSN: 2833-6836; ISBN: 979-8-9870112-4-9. https://asiacall.info/proceedings/index.php/articles/article/view/75

Permana, Y. S. & Cunandar, D. (2025). Guided Note-Taking Strategy in Strengthening Students' Listening Skills. Journal of English Language Pedagogy, Literature, and Culture, Vol. 10 No. 1 2025. STKIP Pancakarya Tasikmalaya

DOI: 10.35974/acuity.v10i1.3777

https://files.eric.ed.gov/fulltext/EJ1460114.pdf

Phuong, H. Y., Lam, T. C. T., & Le, T. T. (2023). Students' perceptions of using note-taking in consecutive interpreting assignments. CTU Journal of Innovation and Sustainable Development, 15(1), 8-21.

https://ctujs.ctu.edu.vn/index.php/ctujs/article/view/547

Suritsky, S.K. (1993). Notetaking difficulties and approaches reported by university students with learning disabilities. Retrieved from: https://files.eric.ed.gov/fulltext/EJ450310.pdf

Tran, T. H. T. (n.d.). Nghiên cứu hành động về việc sử dụng dịch ngữ pháp (Dictogloss) vào dạy kĩ năng nghe cho sinh viên không chuyên tiếng Anh tại Đại Học Quốc Gia Hà Nội. Retrieved from:

https://www.studocu.vn/vn/document/truong-dai-hoc-tra-vinh/tieng-anh-chuyen-nganh-logistics/410187vx22nna5a-tai-lieu-tham-khao-dictogloss/64670667

Tran, T. Q. and Duong, T. M. (2020) Insights into Listening Comprehension Problems: A Case Study in Viet Nam, PASAA: Vol. 59: Iss. 1, Article 4.

DOI: 10.58837/CHULA.PASAA.59.1.4

Available at: https://digital.car.chula.ac.th/pasaa/vol59/iss1/4

Vo, N. T. T. (2026). The impact of effective note - taking techniques on listening comprehension among EFL English majors, Viet Nam Journal of Educational Sciences,

ISSN 2615-8957; e-ISSN 3030-4490 Vol. 22, No. 01 (2026), http://vjes.vnies.edu.vn/en/impact-effective-note-taking-techniques-listening-comprehension-among-efl-english-majors

Appendix

Effective Listening and Note-Taking Checklist for Freshmen

Process	***Techniques***
Before the lecture: *Preparation*	- Preview the lecture topic before class. - Learn important vocabulary related to the lesson. - Read lecture slides or materials if available. - Bring notebooks, pens, laptop, or tablet. - Prepare a simple note-taking format (outline, Cornell notes, mind map). - Sit in a place where you can hear the lecturer clearly. - Get enough rest to improve concentration.
During the lecture: *Listening skills*	- Focus on understanding the main ideas, not every word. - Listen for signal words such as: "First…", "The main point is…", "For example…", "In conclusion…" - Pay attention to repeated ideas and emphasized words. - Ignore small unknown words if the general meaning is clear. - Stay focused and avoid distractions from phones or classmates. - Ask questions if important points are unclear.
Note-taking skills	- Write keywords instead of full sentences. - Use abbreviations and symbols - Record only important ideas and examples. - Separate main ideas from supporting details. - Leave space to add information later. - Use bullet points or numbering to organize notes. - Mark unclear points with "?" for later review. - Do not try to write everything the lecturer says.
After the lecture: *Review and Improvement*	- Review notes - Add missing information while the lecture is still fresh. - Highlight key concepts and important vocabulary. - Rewrite messy notes into a clearer format. - Compare notes with classmates if necessary. - Compare AI-organized notes with their own notes. - Create short summaries of the lecture in your own words. - Practice listening again to difficult parts after class.
Self-reflection checklist	After each lecture, ask yourself: - Did I understand the main ideas? - Did I lose concentration during the lecture? - Which listening problems did I face most? - Did my note-taking method help me?

	- What should I improve for the next lecture?
Golden rules for freshmen	*a. Listen for meaning, not every word.* *b. Write less, understand more.* *c. Review notes quickly after the lecture.* *d. Practice listening every day.* *e. Use technology as support, not dependence.*

#161 BRIDGING THE RHETORICAL GAP: THE CONDITIONAL ROLE OF PERCEIVED DISCURSIVE PROXIMITY IN LEARNING AUTHORIAL VOICE FROM MODEL TEXTS

Tran Huu Thien[1], Pham Huu Duc[2]

Abstract

This study examines how L2 learners perceive metadiscourse as a usable resource for academic writing after exposure to model texts. Although previous research has shown that exposure can help learners notice discourse features, less is known about whether learners see these features as understandable, relevant, and reusable in their own writing. A controlled post-exposure between-subject design was conducted with 190 Vietnamese university students who read one of three model texts differing in metadiscourse visibility. Data included Likert-scale measures of elaborative processing (ELAB), perceived discursive proximity (PDP), perceived metadiscourse uptake (PMU), and open-ended responses. Quantitative findings showed that text condition and PDP positively predicted PMU, whereas ELAB negatively predicted PMU. Qualitative findings suggested that learners evaluated texts through academic authority, accessibility, complexity, and reader guidance. The familiar metadiscourse condition was perceived as the most accessible and reusable. The study reconceptualizes metadiscourse learning as an issue of perceived usability rather than simple exposure and highlights discourse accessibility as an important factor in L2 academic writing.

Keywords: discursive proximity, elaborative processing, L2 academic writing, metadiscourse, model texts, perceived usability

1. Introduction

Academic writing requires writers not only to present information but also to organize arguments, express stance, and guide readers through the development of ideas. These functions are commonly realized through metadiscourse, a set of linguistic resources that enables writers to manage writer-reader interaction and position themselves within disciplinary communities (Hyland, 2005, 2009, 2017). In second language (L2) writing contexts, metadiscourse plays a particularly important role because learners must acquire not only linguistic forms but also the rhetorical conventions through which academic knowledge is communicated and evaluated. To support this process, writing instruction frequently employs model texts. Within genre-based approaches, model texts are used to make rhetorical practices visible by illustrating how writers employ transitions, hedges, self-mentions, engagement markers, and other discourse resources to achieve communicative purposes (Swales, 1990; Tardy, 2006; Yasuda, 2011). Exposure to such texts is often assumed to facilitate learning through observation and imitation. However, the pedagogical value of model texts may depend on more than the visibility of discourse features. Learners may notice metadiscourse forms

[1] *Tran Huu Thien, International Uninversity. VNU – HCMC, email: Email: ENENIU22111@student.hcmiu.edu.vn*

[2] *Pham Huu Duc, International Uninversity. VNU – HCMC, email: phduc@hcmiu.edu.vn*

without fully understanding their rhetorical functions or perceiving them as relevant to their own writing.

This issue reflects a broader shift in metadiscourse research. Earlier frameworks distinguished textual and interpersonal functions of metadiscourse (Crismore & Vande Kopple, 1988), whereas more recent interactional approaches emphasize how metadiscourse functions as a resource for negotiating relationships between writers and readers (Hyland, 2005, 2017). From this perspective, learning metadiscourse involves more than recognizing linguistic markers. It also requires understanding when, why, and under what circumstances those resources can be effectively used. Consequently, the challenge facing L2 learners may not simply be one of exposure but one of perceived usability.

Genre research further highlights this concern. Although model texts can increase awareness of rhetorical resources, learners often experience difficulty transferring those resources into their own writing (Tardy, 2006; Tardy et al., 2020). This difficulty has been linked to conditional knowledge, namely, understanding when and how discourse resources should be applied in new rhetorical situations. As a result, the transition from noticing to meaningful uptake remains insufficiently understood.

The present study addresses this gap by examining learners' perceptions of metadiscourse after exposure to model texts that vary in metadiscourse visibility. Rather than focusing on writing performance, the study investigates how learners evaluate the accessibility, relevance, and potential usability of metadiscourse as a resource for future writing. Three constructs are central to this investigation: Elaborative Processing (ELAB), which captures learners' perceived engagement with model-text discourse; Perceived Discursive Proximity (PDP), which reflects the extent to which academic discourse appears understandable and accessible; and Perceived Metadiscourse Uptake (PMU), which represents learners' perceived readiness to understand and reuse metadiscourse resources.

By examining the relationships among model-text condition, ELAB, PDP, Need for Cognition (NCS), and PMU, this study reconceptualizes metadiscourse learning as an issue of perceived usability rather than simple exposure. In doing so, it contributes to ongoing discussions in L2 writing, genre learning, and metadiscourse research by highlighting the role of discourse accessibility in shaping learners' perceptions of academic writing resources.

2. Literature review

2.1 Metadiscourse in L2 Academic Writing

Academic writing is increasingly viewed as an interactional activity in which writers organize knowledge, express stance, and guide readers toward preferred interpretations. Within this perspective, metadiscourse has become a central concept because it enables writers to manage writer-reader interaction and construct an academic identity (Hyland, 2005, 2009, 2017). Rather than functioning as mere textual markers, metadiscourse resources help writers establish credibility, negotiate commitment, and signal awareness of readers' interpretive needs.

Early frameworks distinguished between textual and interpersonal dimensions of metadiscourse (Crismore & Vande Kopple, 1988). While these classifications provided useful analytical categories, later research emphasized the rhetorical and interactional functions of metadiscourse. Hyland (2005, 2017), for example, argues that features such as transitions, hedges, boosters, self-mentions, and engagement markers should be understood as resources

through which writers position themselves and manage relationships with readers. Consequently, metadiscourse is not simply a matter of textual organization but a means of participating in disciplinary discourse communities.

This interactional perspective is particularly relevant in L2 writing contexts. Learners may recognize metadiscourse forms but still struggle to understand their rhetorical significance or use them appropriately in academic writing. Moreover, metadiscourse conventions vary across disciplines and communicative contexts, making effective use dependent on more than grammatical knowledge alone (Cao & Hu, 2014). The challenge for learners therefore involves not only recognizing discourse resources but also understanding their communicative value.

2.2 Model Texts and Genre-Based Writing Instruction

Model texts occupy a central role in genre-based writing instruction because they make rhetorical practices visible to learners. Swales (1990) argues that genres provide recurring solutions to communicative problems, while model texts illustrate how these solutions are realized through language. Similarly, Yasuda (2011) demonstrates that genre-based instruction helps learners connect linguistic choices with rhetorical purposes, encouraging deeper engagement with academic discourse.

However, genre learning involves more than exposure to exemplary texts. Tardy (2006) proposes that genre acquisition develops through multiple pathways, including instruction, feedback, participation, and textual exposure. Learners may successfully identify rhetorical features in model texts yet encounter difficulties when attempting to transfer those features into their own writing. This challenge reflects the complexity of genre knowledge, which includes not only formal knowledge of textual features but also rhetorical and contextual knowledge (Tardy et al., 2020).

A particularly important component of genre knowledge is conditional knowledge, or knowledge of when and how discourse resources should be applied in new situations. Learners may understand what a hedge or transition does within a model text while remaining uncertain about its appropriate use in their own writing. Consequently, visibility does not automatically translate into usability. Model texts may make discourse resources observable, but learners must still perceive those resources as meaningful and available for future use.

2.3 Elaborative Processing and Its Limits in L2 Writing

The Elaboration Likelihood Model (ELM) offers a useful framework for understanding how learners process model texts. According to Petty and Cacioppo (1986), elaboration refers to careful consideration of message content and argument quality. Higher levels of elaboration are generally associated with deeper cognitive engagement and more stable judgments.

Applied to L2 writing, elaborative processing may indicate attention to rhetorical meaning, argument structure, and discourse organization. Learners who actively analyze how writers construct claims or guide readers may be more likely to notice metadiscourse resources. However, elaboration may also reflect cognitive effort and processing difficulty rather than successful understanding. Self-reported engagement can therefore capture both productive attention and increased processing burden (Petty et al., 2009).

Research on writing motivation and self-regulation further suggests that effort and learning outcomes are not always positively related (Hayes, 2018). Learners may invest substantial cognitive effort because a text is difficult rather than because it is accessible. Consequently, elaborative processing should be interpreted cautiously when examining perceptions of metadiscourse learning.

2.4 Perceived Discursive Proximity as a Usability Construct

To explain why some discourse resources appear more usable than others, this study introduces the construct of Perceived Discursive Proximity (PDP). PDP refers to learners' perceptions that academic discourse is understandable, relevant, and potentially usable for their own writing. The construct focuses on perceived accessibility rather than actual proficiency and seeks to explain how learners evaluate discourse resources after exposure to model texts.

The theoretical foundations of PDP draw on academic literacies, sociocultural theory, and motivational research. Within the academic literacies tradition, literacy is viewed as a socially situated practice involving identity, power, and participation rather than merely the acquisition of skills (Street, 1984; Lea & Street, 1998). Students may possess linguistic knowledge yet still experience difficulty because academic discourse appears distant from their own communicative practices. PDP builds on this insight by focusing on learners' perceptions of access to discourse practices.

Sociocultural theory provides a complementary perspective. Vygotsky (1978) argues that learning occurs through participation in socially mediated activities, while Lantolf and Thorne (2006) emphasize the importance of mediation in language development. From this perspective, model texts function as mediational resources that can either reduce or reinforce perceived distance between learners and academic discourse. PDP therefore reflects learners' judgments regarding the extent to which discourse resources fall within their perceived range of participation.

Motivational theories further suggest that learners are more likely to internalize resources they perceive as personally meaningful and relevant (Ryan & Deci, 2000). Even highly valued discourse practices may fail to support learning if they appear inaccessible or disconnected from learners' identities. PDP captures this evaluative dimension by examining whether academic discourse feels close enough to be appropriated.

Importantly, PDP differs from several related constructs. Unlike self-efficacy, PDP concerns perceived alignment with discourse rather than confidence in performance. A learner may believe they can complete a writing task while simultaneously perceiving the discourse of a model text as distant and difficult to emulate. PDP also differs from perceived difficulty, which focuses primarily on task complexity, and from task value, which concerns usefulness or importance. Instead, PDP specifically addresses learners' perceptions of discourse accessibility and potential usability.

2.5 Previous Studies

Research on L2 writing has increasingly emphasized the roles of identity, motivation, and discourse participation. Casanave (2019) highlights how learners negotiate academic identities, while Lehman (2018) demonstrates the importance of belonging and participation in academic communities. Studies by Dewaele and Dewaele (2018) and Han and Hiver (2018), further illustrate how engagement and learning are shaped by contextual and affective factors.

Despite these developments, relatively little research has examined learners' perceptions of metadiscourse as a usable resource immediately following exposure to model texts. Existing studies have largely focused on noticing, awareness, or textual analysis rather than learners' evaluations of accessibility and potential reuse. Consequently, the relationship between discourse visibility and perceived usability remains insufficiently understood.

2.6 Conceptual Framework

The present study integrates insights from metadiscourse research, genre theory, cognitive processing, and sociocultural perspectives to explain perceived metadiscourse uptake. Model-text condition represents variation in metadiscourse visibility, ranging from minimal to familiar and advanced forms of discourse support. Elaborative Processing (ELAB) reflects learners' perceived engagement with rhetorical meaning, while Perceived Discursive Proximity (PDP) captures discourse accessibility. Need for Cognition (NCS) is included as a dispositional cognitive factor. Perceived Metadiscourse Uptake (PMU) serves as the outcome variable representing learners' perceived readiness to understand and reuse metadiscourse resources.

To examine these relationships, the study employs a conditional process framework. Model-text condition is treated as the independent variable, ELAB as a mediator, PMU as the dependent variable, PDP as the first moderator, and NCS as the second moderator. This framework allows the study to investigate not only whether model texts influence uptake but also how discourse accessibility and cognitive tendencies shape that relationship. Through this perspective, metadiscourse learning is conceptualized as a process of perceived usability rather than simple exposure.

3. Methodology

3.1 Research Design

This study employed a controlled post-exposure mixed-methods design to examine learners' perceptions of metadiscourse after exposure to model texts with different levels of metadiscourse visibility. The design combined quantitative survey data with qualitative open-ended responses, allowing the study to capture both measurable patterns and learners' interpretations of academic discourse (Creswell & Creswell, 2018; Johnson et al., 2007). Rather than examining actual writing performance, the study focused on learners' immediate perceptions following exposure to model texts. Consequently, Perceived Metadiscourse Uptake (PMU) was conceptualized as a perception-based indicator of readiness to use metadiscourse rather than evidence of successful transfer into writing.

3.2 Participants

Participants were 190 Vietnamese undergraduate students enrolled in English-medium courses at a university in Vietnam. All participants had previously completed introductory academic writing instruction and therefore shared comparable levels of familiarity with argumentative writing conventions. Students were randomly assigned to one of three model-text conditions. Random assignment was employed to reduce selection bias and enhance comparability across groups (Mackey & Gass, 2022).

3.3 Materials

Three argumentative model texts were developed specifically for the study. The texts addressed the same topic and followed an equivalent argumentative structure but differed in the visibility of metadiscourse resources. Text 1 contained minimal explicit metadiscourse, reducing overt stance and engagement markers. Text 2 incorporated familiar and highly visible features such as transitions, self-reference, and reader-oriented guidance. Text 3 employed more advanced and implicit forms of metadiscourse, including hedging and indirect positioning strategies.

The design of the materials was informed by genre-based writing research, which emphasizes the pedagogical value of model texts in making rhetorical practices visible (Swales, 1990; Tardy, 2006; Yasuda, 2011). By manipulating metadiscourse visibility while holding

topic and structure constant, the study sought to isolate the influence of discourse accessibility on learners' perceptions.

3.4 Instruments

Four perception-based constructs were measured through Likert-scale questionnaires. Elaborative Processing (ELAB) was assessed using six items adapted from the Elaboration Likelihood Model and measured learners' perceived engagement with rhetorical meaning and textual organization (Petty & Cacioppo, 1986; Petty et al., 2009). Perceived Discursive Proximity (PDP) was measured through seven items developed for this study to assess learners' perceptions of discourse accessibility, interpretability, and potential usability. The construct was theoretically informed by academic literacies, sociocultural theory, and self-determination theory (Lea & Street, 1998; Ryan & Deci, 2000; Vygotsky, 1978).

Perceived Metadiscourse Uptake (PMU) assessed learners' beliefs regarding their ability to understand and potentially reuse metadiscourse resources in future writing. Finally, Need for Cognition (NCS) measured participants' dispositional tendency to engage in effortful cognitive activity and was included as an individual-difference variable (Petty et al., 2009). In addition, participants completed several open-ended questions concerning memorable textual features, perceived academic quality, reader guidance, and the usability of discourse resources. These responses provided qualitative insights into learners' interpretations of the model texts.

3.5 Data Analysis

Data analysis proceeded in several stages. First, reliability and construct validity were assessed through Cronbach's alpha, McDonald's omega, and Confirmatory Factor Analysis (CFA). CFA was conducted to evaluate the measurement properties of ELAB, PDP, PMU, and NCS before proceeding to subsequent analyses (Brown, 2015; Kline, 2023).

Second, multiple regression analyses were performed to examine relationships among the study variables. Model-text condition, ELAB, PDP, and NCS were entered as predictors, while PMU served as the outcome variable. This analysis allowed the study to evaluate the relative contribution of discourse accessibility, cognitive engagement, and learner characteristics to perceived metadiscourse uptake (Hayes, 2014).

Third, a moderated mediation analysis was conducted using PROCESS Model 63 (Hayes, 2014). In this model, model-text condition was specified as the independent variable, ELAB as the mediator, and PMU as the dependent variable. PDP functioned as the first moderator and NCS as the second moderator. This framework enabled the study to investigate whether the indirect effect of model-text condition on perceived uptake through elaborative processing varied according to learners' perceptions of discourse accessibility and their tendency toward cognitive engagement.

Finally, qualitative responses were analyzed using thematic coding procedures (Braun & Clarke, 2006, 2022). Coding focused on recurring themes related to academic authority, accessibility, complexity, reader guidance, and discourse usability. To enhance trustworthiness, coding decisions were documented through an audit trail and supported by representative participant excerpts.

4. Results

4.1 Quantitative Results

A Confirmatory Factor Analysis (CFA) was conducted to evaluate the measurement properties of the four constructs. Although several fit indices suggested acceptable

approximation (χ^2/df = 2.129; RMSEA = .077), the overall measurement model demonstrated only moderate fit (CFI = .820; TLI = .798; SRMR = .087). Consequently, subsequent analyses were conducted using observed composite scores rather than latent variables.

Table 1 presents the results of the multiple regression and conditional process analyses. Model-text condition significantly predicted Perceived Metadiscourse Uptake (PMU) (b = .473, p < .001), indicating that learners exposed to different model texts reported different levels of perceived uptake. Perceived Discursive Proximity (PDP) also emerged as a strong positive predictor of PMU (b = .486, p < .001), suggesting that learners who perceived academic discourse as more understandable and accessible were more likely to report readiness to reuse metadiscourse resources.

In contrast, Elaborative Processing (ELAB) negatively predicted PMU (b = −.451, p < .001). Rather than indicating successful engagement, higher levels of self-reported elaboration may reflect increased processing effort or difficulty when interpreting model-text discourse. Need for Cognition (NCS) showed a smaller but significant positive effect on PMU (b = .154, p = .019). The interaction between model-text condition and PDP was also significant (b = −.252, p < .001), suggesting that the influence of discourse accessibility varied across different text conditions. However, interactions involving ELAB and NCS did not reach statistical significance.

Table 1

Measurement and Regression Results

Analysis / Predictor	**b / Value**	**SE**	**t**	**p**	**95% CI**
Measurement model					
χ^2(df)	523.750(246)	—	—	< 0.001	—
χ^2/df	2.129	—	—	—	—
CFI / TLI / RMSEA / SRMR	0.820 / 0.798 / 0.077 / 0.087	—	—	—	—
ELAB model					

Text	−0.167	0.056	−2.961	0.003	[−0.278, −0.055]
PDP	0.504	0.035	14.245	< 0.001	[0.434, 0.574]
NCS	0.073	0.052	1.400	0.163	—
Text × PDP	−0.029	0.048	−0.605	0.545	—
Text × NCS	−0.191	0.072	−2.631	0.009	[−0.334, −0.047]
PMU model					
Text	0.473	0.071	6.611	< 0.001	[0.332, 0.614]
ELAB	−0.451	0.096	−4.686	< 0.001	[−0.641, −0.261]
PDP	0.486	0.063	7.662	< 0.001	[0.361, 0.611]
NCS	0.154	0.065	2.351	0.019	[0.024, 0.284]
Text × PDP	−0.252	0.059	−4.243	< 0.001	[−0.369, −0.135]
Text × NCS	−0.021	0.091	−0.229	0.818	—

ELAB × PDP	−0.010	0.043	−0.229	0.818	—

4.2 Qualitative Results

Analysis of the open-ended responses identified three recurring dimensions shaping learners' evaluations of the model texts: academic authority, argumentative manageability, and reader guidance. Participants frequently associated academic quality with evidence-based reasoning, formal language, and logical organization.

Distinct perceptual profiles emerged across the three text conditions. Text 1 was commonly described as formal and objective but relatively distant. Text 2 was viewed as the most accessible, with participants highlighting explicit transitions, reader-oriented guidance, and clear argumentative structure. Text 3 was perceived as academically sophisticated but also more demanding due to its implicit rhetorical features and greater discourse complexity (Table 2).

Table 2

Qualitative Profiles Across Conditions

Dimension	**Text 1**	**Text 2**	**Text 3**
Academic authority	Formal, objective, impersonal	Formal but accessible	Formal, sophisticated, abstract
Accessibility	Moderate to low	Highest accessibility	Lower accessibility due to complexity
Complexity	Academic distance and density	Lowest complexity profile	Highest complexity profile
Reader guidance	Mainly indirect structure	Explicit structure and visible stance	Structural guidance without direct address
Engagement	Limited direct engagement	Most directly engaging	Impersonal but rhetorically guided
Overall profile	Academic distance	Accessible instructional bridge	Advanced academic sophistication

4.3 Integrated Results

The quantitative and qualitative findings converged around the concept of discourse usability. Quantitative analyses demonstrated that perceived accessibility strongly predicted

perceived uptake, while qualitative responses revealed that learners consistently evaluated texts in terms of their usability rather than simply their academic quality.

Text 2 emerged as the most positively evaluated condition because it balanced academic credibility with accessibility. Across both data sources, learners appeared more willing to adopt discourse resources when those resources were visible, understandable, and manageable (Table 3).

Table 3

Integrated Findings

Data source	**Main finding**	**Interpretation for perceived uptake**
Measurement model	Marginal fit: $\chi^2/df = 2.129$, RMSEA = 0.077, CFI = 0.820, SRMR = 0.087	Constructs were interpreted as observed composites.
Quantitative model	Text positively predicted PMU, b = 0.473, $p < 0.001$	Model-text condition was associated with perceived uptake.
Quantitative model	PDP positively predicted PMU, b = 0.486, $p < 0.001$	Discourse access was central to perceived uptake.
Quantitative model	ELAB negatively predicted PMU, b = −0.451, $p < 0.001$	Reported elaboration indexed processing effort as well as engagement.
Quantitative model	Text × PDP was significant, b = −0.252, $p < 0.001$	Text condition mattered less when PDP was high.
Open-ended responses	Text 1 was formal, neutral, and distant	Academic authority was paired with lower perceived closeness.
Open-ended responses	Text 2 was accessible, structured, and reader-facing	Familiar metadiscourse supported perceived usability.

Open-ended responses	Text 3 was sophisticated, abstract, and demanding	Advanced discourse signalled quality but increased perceived distance for some learners.
Integrated pattern	Learners evaluated texts through authority, manageability, and guidance	Perceived uptake depended on how discourse was experienced as usable.

5. Discussion

5.1 Metadiscourse as Usable Discourse

The findings suggest that metadiscourse learning should be understood not only as a matter of exposure but also as a matter of perceived usability. Although model texts have long been used to make rhetorical practices visible in writing instruction (Swales, 1990; Yasuda, 2011), visibility alone does not guarantee uptake. Learners may recognize discourse features without perceiving them as relevant or available for their own writing.

The positive effect of model-text condition on Perceived Metadiscourse Uptake (PMU) supports this argument. Participants exposed to different model texts reported different levels of perceived uptake, indicating that metadiscourse visibility influences how learners evaluate discourse resources. In particular, the familiar metadiscourse condition generated the most favorable responses across both quantitative and qualitative measures. Learners repeatedly described this condition as clear, organized, and reader-friendly, suggesting that discourse resources became more usable when rhetorical functions were made explicit.

This finding extends interactional approaches to metadiscourse (Hyland, 2005, 2017) by shifting attention from discourse production to discourse accessibility. Previous studies have primarily examined how expert writers employ metadiscourse to construct stance and engage readers. The present study instead focuses on how learners evaluate those resources before attempting to use them. From this perspective, metadiscourse learning involves an intermediate stage in which discourse resources become perceived as understandable, relevant, and potentially reusable.

The results also align with genre-based perspectives on writing development. Tardy (2006) and Tardy et al. (2020) argue that genre learning requires more than awareness of textual features. Learners must develop conditional knowledge regarding when and how rhetorical resources can be applied. The present findings suggest that perceived usability may represent an important precursor to this process. Before learners can appropriate discourse resources, they must first perceive those resources as accessible.

5.2 Reinterpreting Elaborative Processing

One of the most unexpected findings was the negative relationship between Elaborative Processing (ELAB) and PMU. Within the Elaboration Likelihood Model, elaboration is typically associated with deeper engagement and more stable judgments (Petty & Cacioppo, 1986). A straightforward prediction would therefore suggest a positive relationship between elaboration and perceived uptake. However, the present findings indicate the opposite pattern.

A possible explanation is that self-reported elaboration captures not only engagement but also processing burden. In L2 academic writing contexts, learners frequently encounter unfamiliar discourse practices, disciplinary expectations, and rhetorical conventions. Greater elaboration may therefore reflect increased effort required to understand a text rather than successful integration of its discourse resources.

This interpretation is consistent with research on writing self-regulation and cognitive processing (Hayes, 2018). Cognitive effort does not always indicate effective learning. Learners may invest substantial mental resources because discourse appears difficult, distant, or unfamiliar. Under such circumstances, high elaboration may signal struggle rather than mastery.

The qualitative findings provide further support for this interpretation. Participants frequently described Text 3 as sophisticated and academically advanced, yet they also viewed it as difficult to understand and emulate. In contrast, Text 2 received lower descriptions of difficulty while simultaneously producing higher levels of perceived uptake. These patterns suggest that accessibility may be more important than processing intensity when learners evaluate discourse usability.

Consequently, elaborative processing should not automatically be interpreted as evidence of successful rhetorical engagement. Instead, it may function as an ambiguous indicator reflecting both cognitive involvement and processing difficulty. Future research should therefore distinguish between productive elaboration and effort associated with discourse complexity.

5.3 Perceived Discursive Proximity as a Condition for Uptake

Among all predictors examined in the study, Perceived Discursive Proximity emerged as one of the strongest predictors of perceived metadiscourse uptake. This finding supports the central theoretical claim of the study: learners are more likely to adopt discourse resources when those resources appear close enough to their existing communicative repertoire.

The concept of PDP builds upon academic literacies research, which emphasizes access to discourse practices rather than simple acquisition of textual skills (Lea & Street, 1998; Street, 1984). From this perspective, learners' difficulties often stem not from a lack of linguistic knowledge but from uncertainty regarding participation in academic discourse communities. PDP operationalizes this issue by focusing on learners' perceptions of discourse accessibility.

The findings also resonate with sociocultural perspectives on learning. According to Vygotsky (1978), learning occurs through participation in socially mediated activities. Model texts can therefore function as mediational resources that either reduce or reinforce perceived distance between learners and academic discourse. When discourse resources appear understandable and manageable, learners may perceive them as available for future use. When discourse appears distant, uptake becomes less likely.

Furthermore, PDP complements motivational perspectives on internalization (Ryan & Deci, 2000). Learners are more likely to adopt resources they perceive as personally meaningful and attainable. The strong relationship between PDP and PMU suggests that discourse accessibility may play a critical role in transforming visible discourse resources into potentially usable ones.

Taken together, the quantitative and qualitative findings indicate that perceived accessibility is not a peripheral factor but a central condition for metadiscourse uptake. This

contribution extends existing metadiscourse and genre-learning research by highlighting the importance of learners' perceptions of discourse availability.

5.4 Pedagogical Implications

The findings have several implications for academic writing instruction. First, model texts should be selected not only for authenticity but also for their ability to make rhetorical functions visible and accessible. Texts that balance academic sophistication with reader-oriented guidance may be particularly beneficial for novice writers.

Second, instruction should move beyond identifying metadiscourse features toward explaining their rhetorical purposes. Helping learners understand why writers use particular discourse resources may increase perceived accessibility and support future transfer.

Finally, writing pedagogy should recognize that learners' perceptions of discourse accessibility can influence uptake. Creating opportunities for guided discussion, reflection, and adaptation of model-text discourse may help learners perceive academic writing as something they can realistically use rather than simply admire from a distance.

6. Conclusion

This study examined how Vietnamese university students perceived metadiscourse after exposure to model texts with varying levels of metadiscourse visibility. By integrating quantitative and qualitative evidence, the findings demonstrated that model-text condition and Perceived Discursive Proximity (PDP) positively predicted Perceived Metadiscourse Uptake (PMU), whereas Elaborative Processing (ELAB) showed a negative relationship with uptake. Qualitative analyses further revealed that learners evaluated model texts in terms of academic authority, accessibility, complexity, and reader guidance, with the familiar metadiscourse condition emerging as the most accessible and reusable.

The study contributes to L2 writing research by reconceptualizing metadiscourse learning as a matter of perceived usability rather than simple exposure. While previous research has focused primarily on noticing discourse features, the present findings suggest that learners are more likely to adopt rhetorical resources when those resources are perceived as understandable, relevant, and accessible. In this regard, PDP offers a useful framework for understanding how learners evaluate their access to academic discourse and how such evaluations shape perceived uptake.

Pedagogically, the findings highlight the importance of selecting model texts that balance academic sophistication with discourse accessibility. Writing instruction should therefore emphasize not only the identification of metadiscourse features but also their rhetorical functions and practical applicability. Future research should examine whether perceived usability predicts actual transfer into writing through longitudinal designs and analyses of learners' written production.

References

Alzubi, A. A. F., & Nazim, M. (2024). Students' intrinsic motivation in EFL academic writing: Topic-based interest in focus. *Heliyon, 10*(1), e24169. https://doi.org/10.1016/j.heliyon.2024.e24169

Berkenkotter, C., Huckin, T. N., & Ackerman, J. (1988). Conventions, conversations, and the writer: Case study of a student in a rhetoric Ph. D. program. *Research in the Teaching of English,* 9–44. https://www.jstor.org/stable/40171130

Braun, V. and Clarke, V. (2006) Using Thematic Analysis in Psychology. Qualitative Research in Psychology, *3*, 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V. and Clarke, V. (2022) *Thematic Analysis: A Practical Guide.* SAGE Publications.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *Journal of Pragmatics, 66,* 15–31. https://doi.org/10.1016/j.pragma.2014.02.007

Casanave, C. P. (2019). Performing expertise in doctoral dissertations: Thoughts on a fundamental dilemma facing doctoral students and their supervisors. *Journal of Second Language Writing, 43,* 57–62. https://doi.org/10.1016/j.jslw.2018.02.005

Creswell, J.W. & Creswell, J.D. (2018) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* Sage, Los Angeles.

Crismore, A., & Vande Kopple, W. J. (1988). Readers' learning from prose: The effects of hedges. *Written Communication, 5*(2), 184–202. https://doi.org/10.1177/0741088388005002004

Dewaele, J.-M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the FL classroom. *Journal of the European Second Language Association, 2*(1), 24–37. https://doi.org/10.22599/jesla.37

Dong, L. (2026). Is it beneficial to strive for perfection in writing? Exploring the relationship between perfectionism, motivational regulation, and second language (L2) writing performance. *Assessing Writing, 67,* 101012. https://doi.org/10.1016/j.asw.2025.101012

Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing, 40,* 44–59. https://doi.org/10.1016/j.jslw.2018.03.001

Hao, X., & Lu, H. (2024). The role of self-efficacy and learner engagement in the relationship between prior achievement and final achievement in EFL writing. *Learning and Motivation, 87,* 102007. https://doi.org/10.1016/j.lmot.2024.102007

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hayes, J. R., & Berninger, V. W. (2014). Cognitive processes in writing. In *Oxford University Press eBooks* (pp. 3–15). https://doi.org/10.1093/acprof:oso/9780199827282.003.0001

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.

Hyland, K. (2009). *Academic discourse: English in a global context.* Continuum International Publishing Group.

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113,* 16–29. https://doi.org/10.1016/j.pragma.2017.03.007

Hyland, K., & Sancho Guinda, C. (Eds.). (2012). *Stance and voice in written academic genres.* Palgrave Macmill.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112–133. https://doi.org/10.1177/1558689806298224

Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling.* Guilford Publications.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development.* Oxford University Press.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education, 23(*2), 157–172. https://doi.org/10.1080/03075079812331380364

Lehman, I. M. (2018). *Authorial presence in English academic texts: A comparative study of student writing across cultures and disciplines.* Peter Lang. https://doi.org/10.3726/b14175

Mackey, A., & Gass, S. M. (2022). *Second language research: Methodology and design* (3rd ed.). Routledge.

Petty, R. E., Barden, J., & Wheeler, S. C. (2009). The elaboration likelihood model of persuasion: Developing health promotions for sustained behavioral change. In R. J. DiClemente, R. A. Crosby, & M. C. Kegler (Eds.), *Emerging theories in health promotion practice and research* (2nd ed., pp. 185–214). Jossey-Bass/Wiley.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). Academic Press.

Philippakos, Z. A. T., MacArthur, C., & Wang, C. (2023). Writing motivation of college students in basic writing and first-year composition classes: Confirmatory factor analysis of scales on goals, self-efficacy, beliefs, and affect. *Journal of Learning Disabilities, 56*(1), 72–92. https://doi.org/10.1177/00222194211053238

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Sengupta, K. (1999). *Choice rules with fuzzy preferences: Some characterizations. Social Choice and Welfare, 16*(2), 259–272. https://doi.org/10.1007/s003550050143

Street, B. V. (1984). *Literacy in theory and practice*. Cambridge University Press.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Tardy, C. M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of Second Language Writing, 15*(2), 79–101. https://doi.org/10.1016/j.jslw.2006.04.003

Tardy, C. M., Sommer-Farias, B., & Gevers, J. (2020). Teaching and researching genre knowledge: Toward an enhanced theoretical framework. *Written Communication, 37*(3), 287–321. https://doi.org/10.1177/0741088320916554

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*(2), 91–112.

VanPatten, B., & Williams, J. (Eds.). (2015). *Theories in second language acquisition: An introduction* (2nd ed.). Routledge.

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wang, X., Hussin, S., & Aziz, A. A. (2025). Investigating the relationship between self-regulated writing strategies and academic writing performance. *Arab World English Journal, 16*(3), 345–358. https://doi.org/10.24093/awej/vol16no3.21

Yasuda, S. (2011). Genre-based tasks in foreign language writing: Developing writers' genre awareness, linguistic knowledge, and writing competence. *Journal of Second Language Writing, 20*(2), 111–133. https://doi.org/10.1016/j.jslw.2011.03.001 .

Zhao, F., & Chi, R. (2026). Identity positioning in teaching intercultural communication: A case study of Chinese EFL teachers. *Intercultural Communication Education, 9*(1), 103267. https://doi.org/10.29140/ice.v9n1.103267

#164 INTEGRATING AI LITERACY AND DELIVERY INSTRUCTION IN PUBLIC SPEAKING CLASSES FOR VIETNAMESE UNDERGRADUATE STUDENTS

Mai Thị Mỹ Hạnh[1]

Abstract

The rapid integration of generative AI tools into the academic learning process has fundamentally transformed traditional public speaking instructions. When undergraduate students can easily produce speech scripts and slides with AI assistance, the instructors face two pressing questions: how do teachers encourage authentic oral communication, and how does delivery training impact student performance when presenting AI-assisted content? To address these issues, this study employs a reflective teaching approach to evaluate instructional practices that balance AI literacy with intensive oral delivery instruction. This dual-framework instructional model was implemented across two Public Speaking courses with 80 Vietnamese undergraduate English majors over one academic semester. In the first phase, students were taught prompt engineering to critically evaluate, refine, and personalize AI-generated outlines; and then came the delivery instruction phase, where structured training in vocal variety, pausing, and intonation control were provided. Data were collected from classroom observations, teacher reflections, and student feedback and performance results. Findings show that systematic delivery training produced observable improvements in students' oral presentation performance, while AI literacy activities enhanced personal ownership of the language. The paper provides practical recommendations for TESOL practitioners teaching oral communication in AI-integrated language teaching contexts.

***Keywords*:** AI literacy, public speaking, oral delivery, EFL/ESL, TESOL

1. Introduction

We are at the time when the integration of GenAi tools has completely transformed and reshaped English Language Teaching (ELT), particularly with the subject of Public Speaking. The AI-driven tools can now help undergraduate students create logical and grammatically correct speech scripts in a matter of minutes. Although this has raised concerns regarding academic integrity and a potential threat to erosion of critical thinking, it also provides Public Speaking instructors an opportunity to re-evaluate what is important in teaching this subject. In public speaking, a good script is not equal to a good presentation; it depends on other features, such as vocal variety, pausing, and intonation, facial expressions and body language. When students passively copy AI-generated text, their speeches often result in flat, unengaging, and poor performances, since the language used in the text tends to exceed students' linguistic competence. To address this challenge, this study proposes an instructional framework implemented in an undergraduate public speaking course for English majors at a university in Vietnam. Previous AI studies have mainly focused on writing, academic integrity, or language accuracy, this study addresses how AI-generated content can be transformed into personally owned oral performance through delivery training. Rather than banning the use of AI tools, this study investigates an alternative instructional method. Specifically, it examines how

[1]*MA., Faculty of Foreign Languages, HCM City University of Foreign Languages & Information Technology (HUFLIT), Email: hanhmtm@huflit.eud.vn*

embedding AI literacy and prompt engineering into the brainstorming and outline phrases of speech preparation, combined with intensive delivering training, supports students' public speaking development. The public speaking classroom, through this dual instructional framework model, demands that students inhabit the language, use it as their own to say something that really matters to them in a confident manner. Specifically, the study aims to assess the effectiveness of targeted training in prosodic delivery features, specifically vocal variety, strategic pausing, and intonation control, in improving students' presentation performance; and analyze the influence of the integration of prompt engineering and AI literacy activities into a public speaking course on students' sense of language ownership and authentic engagement with AI-generated content.

2. Literature Review

Artificial intelligence (AI) in language learning and teaching

Studies have increasingly reported positive influences of the pedagogical potential of integrating AI in language learning on learners' oral performance, speaking confidence, and interactional competence. AI tools can generate different argument structures and language styles, providing students with examples they can emulate, thereby boosting confidence and broadening their linguistic repertoire (Steiss et al., 2024). A great amount of scholarly response focused on writing contexts - examining how AI tools affect the writing process, what they mean for academic integrity, and how writing pedagogy might adapt to a landscape in which fluent, coherent text can be generated with minimal human effort (DuBose et al.,2023). This is understandable given that text generation is the most visible and disruptive capability of AI tools, but it has left a considerable gap in the literature: the implications of AI for oral communication instruction, particularly in public speaking skills, remain relatively underexplored. With the aid of AI-driven tools, students can generate the content of their speeches or their presentations before delivering them to a human audience; and this raises questions about authenticity and the purpose of oral communication instruction. The challenges this practice creates are what several scholars have described as a *competence-performance gap* related to AI-assisted language production (Gilea et al., 2025). When AI generates sophisticated language that exceeds learners' proficiency levels, a visible mismatch emerges between the fluency of the text and the speaker's ability to elaborate on or defend it in real-time communication. This mismatch is particularly notable in public speaking contexts. Addressing this gap is one of the central motivations for the dual-framework model described in this paper.

Authenticity and Speaker's voice in Public Speaking

Research on authenticity in EFL public speaking has established that audiences, including instructors and peers, are sensitive to markers of speaker ownership that go beyond linguistic fluency. Joughin (2010), in a study of oral assessment in higher education, stated that evaluators intuitively recognize students who appeared to *know* their material and those who appeared to be *reproducing* it, and that this distinction shaped evaluative judgments even when assessment criteria did not name it. More recently, studies of presentation skills in EFL university contexts have noted that the highly rated student presenters tend to be those who integrate personal experience and locally relevant examples into their talks - precisely the behavior observed and encouraged in the present study (Chen, 2019). These findings prove that rhetorical ownership is not a desirable quality but an observable and assessor-recognizable feature of effective oral performance.

Delivery Instruction and Oral Performance in Public Speaking

Public speaking competence emphasizes oral delivery, together with content preparation and organization. Delivery competence is the ability to present within a specific context, including the capacity to evoke emotions and regulate one's behavior (Zhang et al., 2019). A speaker with delivery competence knows how to communicate non-verbally, using eye contact, gestures, and voice to convey their message to the audience (Lucas & Stob, 2020). Delivery features such as vocal variety, eye contact, fluency, and body language have been identified as significant predictors of audience perception of speaker credibility, engagement, and overall effectiveness. Studies have shown that ESL/EFL speakers often struggle with prosodic features, specifically intonation, pausing, and stress, in ways that can impede communicative clarity regardless of content quality (Derwing & Munro, 2005; Gilbert, 2008). This shows that it is imperative to treat delivery not as a concern but as a core instructional target in public speaking courses for ESL/EFL learners.

Intonation control is perhaps the most complex instructional challenge of the three delivery features targeted in this study, both because of its linguistic complexity and because of the specific transfer challenges it has for Vietnamese-speaking learners. Vietnamese, as a tonal language, uses pitch variation at the syllable level to mark lexical meaning, a fundamentally different function that draws on some of the same phonological resources but deploys them in ways that can interfere with the acquisition of English intonation patterns. Research on Vietnamese learners' English prosody has documented consistent patterns of L1 transfer in intonation production, including a tendency toward flatter overall contours, reduced use of the high-rising terminal, and syllable-timed rhythm that can obscure the stress-timed information structure of English (Tran, 2006). Targeted intonation instruction, especially when coupled with activities and models of native or proficient speaker norms, has demonstrated positive effects on both production accuracy and learner confidence (Nguyen, 2004).

Empirical studies that have incorporated explicit delivery instruction into public speaking or oral presentation courses report consistently positive findings. King (2002) stated that iterative practice combined with instructor modeling of delivery behaviors produced benefits not only in technical delivery but also in student confidence and willingness to communicate. In TESOL contexts specifically, delivery-focused instruction has been linked to improvements in comprehensibility ratings by both teacher assessors and peer listeners (Derwing et al., 2004). Together, these three strands establish the conceptual foundation for the dual-framework model and situate the study's research questions.

Although significant improvements have been made in AI-supported language education, research gaps remain, specifically concerning AI's impact on public speaking instruction for EFL/ESL learners in Vietnamese contexts. There is limited theoretical and empirical research exploring how AI tools affect presentation performance. This highlights an area in urgent need of further research.

The dual-framework Instructional Model

In the era when AI can assist students to produce polished speech scripts within seconds, the public speaking classrooms then become a performance in which students simply read aloud what AI has written for them, the public speaking class therefore may lose its main purpose, which is producing confident speakers who can connect with an audience in their own voice. The dual-framework instructional design was developed in response to this challenge.

Instead of banning AI use, which proves impractical in the contemporary classrooms, the model integrates it into the course design through two parallel instructional tracks. *Phase 1* - AI Literacy - addresses the content of the speeches by guiding students to analyze,

personalize, and engage with AI-generated drafts. *Phase 2* - Delivery Instruction - provides skill-focused training in the main features of public speaking that distinguish an interesting speaker from a script reader, regarding vocal variety, pausing, and intonation control. The model was implemented across two Public Speaking courses enrolling a total of 80 Vietnamese undergraduate English majors over one semester. Both courses followed the same instructional sequence, allowing the instructor to observe patterns across class groups and refine activities in response to student needs.

Phase 1 - AI Literacy Instruction

The AI Literacy phase was designed around a single guiding principle: students should be the authors of their speeches, even when AI assists in the first draft. Instead of using AI-generated text as a finished product to be memorized and delivered, this phase treated it as raw materials, the first draft that required the student's own knowledge, experience, and communicative intention to make it become their own speech.

Key Activity: Revision and personalization of AI-generated drafts

The primary instructional activity in this phase involved a revision process applied to AI-generated speech drafts. Students were first taught the sample prompts and were invited to use an AI tool of their choice (ChatGPT or Gemini) to generate a draft/outline on their assigned or chosen speech topic(s). This draft served as a starting point rather than a final text. Students were then guided through a process in which they were asked to identify the text / passages that did not reflect their own knowledge or experience, replace generic AI examples with personal stories or locally relevant examples/illustrations, adjust the formality and tone to match their own natural speaking style, and insert their own linking phrases and signposting rather than using the AI's structural scaffolding. This revision process was not presented as a corrective measure against AI use, but as a standard part of the speech writing process - the same critical stage that a skilled writer would apply to any draft. The instructional message emphasized that AI could help you structure your ideas, but only *you* can make this speech worth listening to. More importantly, this activity served as preparation for the oral performance itself.

Assessment of AI Literacy outcomes

AI literacy was assessed through the content section of the course's rubric. Content scores were based on students' personal engagement with and investment in their material during the presentation - specifically, their ability to analyze and criticize the AI-script, respond to audience or instructor prompts, and integrate personal examples or locally based evidence.

Phase 2 - Delivery Instruction

While the AI Literacy phase addressed what students said, the Delivery Instruction phase focused on how they said it. A polished script read in a flat monotone communicates far less than a modest script delivered with confidence and rhythm. For EFL learners particularly, whose L1 prosodic patterns (Vietnamese) may transfer in ways that affect intelligibility and audience engagement in English, repeated delivery training is not supplementary, but it is crucial. The delivery instruction sequence focused on three main features: vocal variety (modulating pitch, pace, and volume to catch audience attention and emphasize key ideas), strategic pausing (using silence to signal transitions, allow audience time to think, and project speaker confidence), and intonation (managing rising and falling patterns to convey meaning and attitude in English).

Activity 1: Drilling Exercises

Drilling activities formed a cornerstone of the delivery instruction and was integrated into in-class sessions throughout the semester. Drilling tasks targeted each of the three delivery features in focused practice contexts before students were asked to apply them in full presentations. For vocal variety, students practiced reading short passages aloud with instructor-guided cues. As for pausing, students worked with marked-up scripts in which pause points were indicated, practicing the deliberate act of stopping, and holding the silence before continuing next points. Many students initially found this uncomfortable, as silence in performance contexts often considered as error or not fluent in their prior learning experience. Repeated drilling helped normalize the pause as a tool rather than a failure. For intonation and stress, students practiced contrastive stress patterns, fall-rise sequences for hedging and suggestion, and the high-rising that signals incompleteness in English, the features that differ from Vietnamese tonal patterns and require attention.

Activity 2: Feedback on Recorded and Live Presentations

Feedback sessions were structured around the same delivery rubric used for formal assessment, training students to observe and articulate delivery features in each other's performances. By using the same rubric for both peer feedback and formal assessment, students developed metalinguistic awareness of delivery features while simultaneously learning to evaluate their peers' performances. Students were grouped to listen to short-recorded practice presentations or a live speech and provide written feedback using a guided observation form that prompted specific comments on vocal variety, pausing, and intonation. This activity served as a dual function. As feedback providers, students developed an awareness of delivery features that sharpened their own self-monitoring during practice. As feedback recipients, students received peer observations that often carried different power than instructor feedback - hearing from a classmate that "*your voice went flat in the middle section*" proved, in many cases, more helpful and impactful than the same note from a teacher. Live peer feedback sessions were also conducted in class, with small groups observing a presenter and offering real-time observations immediately following each practice round.

3. Methodology

This study adopted a reflective classroom research design to examine the effectiveness of integrating AI literacy and oral delivery instruction into the Public Speaking course for second-year students majoring in English. This study investigates a pedagogical intervention implemented by the course instructor in a real classroom context; therefore, the design was chosen to allow for pedagogical adjustments during implementation while evaluating the impact on student performance and language ownership. This reflective classroom research can be more appropriate than experimental or quasi-experimental design for this stage of inquiry. The study can be framed as exploratory practitioner research aiming to generate pedagogical insights in public speaking courses. The course instructor acted as a practitioner-researcher, collecting and interpreting evidence from direct classroom experience, self-reflection, and student interactions throughout one academic semester.

Participants were 80 second-year English-major students in two Public Speaking classes, which shared similar learning conditions, at the Faculty of Foreign Languages, HUFLIT. All students, at intermediate to upper-intermediate English proficiency (B1–B2 CEFR level), had prior experience with basic presentation tasks but limited systematic training in delivery skills or AI-assisted content creation.

Data were collected from three sources. Accordingly, *Classroom observations* were conducted throughout the semester. The instructor maintained notes about instructional

practices, use of AI-generated versus personalized content, student engagement patterns, and difficulties (if any) in both the AI literacy and presentation skills training activities. *Teacher reflections* were recorded systematically across the semester in the form of a reflective journal and teaching logs, maintained by the course instructor. These entries captured pedagogical decisions, student responses, and adjustments made to instruction. In addition, *Student feedback and performance results* were collected at specific assessment points in the semester. There were 3 in-class mini speeches (3-5 minutes) and 1 video-recorded speech. All presentations were scored using a validated analytic rubric (adapted from Lucas, 2015, and institution-specific criteria) assessing both content ownership and delivery effectiveness. The rubric included subscales for vocal delivery (25%), non-verbal delivery (20%), content adaptation/personalization (25%), organization (15%), and audience engagement (15%). Student feedback was also collected via informal written reflections on their experience of integrating AI tools into the speech preparation process. Classroom observation notes and teacher reflections were analyzed thematically, following an inductive coding procedure to identify recurring patterns related to student engagement, challenges, and instructional effectiveness. Student performance data were examined to assess change in delivery-related scores across the semester, while written feedback was analyzed for themes pertaining to students' perceived agency, language ownership, and attitudes toward AI-assisted speech preparation. Triangulating these data sources allowed the study to develop a multi-angled understanding of classroom dynamics and to propose evidence-based adjustments for desired learning outcomes.

The course instructor acted as a practitioner-researcher, collecting and interpreting evidence of student learning outcomes and engagement throughout one academic semester of 12 weeks. Since potential biases might arise from this dual role, different measures were implemented to ensure objectivity in data collection, including student feedback and informal discussions with students during break time for each teaching session.

This methodology enabled a comprehensive examination of how AI literacy and delivery instruction can be effectively combined to promote oral communication skills in an AI-rich educational environment.

4. Results

The findings are organized around three interrelated themes that emerged across the three data sources, including classroom observations, teacher reflections, and student feedback and performance results. The first theme addresses the impact of structured delivery training on students' oral presentation performance. The second examines how students engaged with and took ownership of their AI-assisted content. The third explores the productive interplay between the two instructional tracks, showing how gains in one dimension frequently amplified and made visible gains in the other.

Impact of Delivery training on Oral performance

The most significant and measurable finding across the semester was that public speaking delivery training produced considerable improvements in students' presentation performance across all three features: vocal variety, strategic pausing, and intonation control. Improvements were evenly distributed, proving that the drilling and feedback activities benefited students across the range of prosodic features addressed in the course.

Vocal variety: At the beginning of the semester, classroom observations showed that students delivered their speeches in a flat pitch range, with pace and volume remaining flat from the opening to the end of the speech. Teacher reflections from the first 3 weeks noted that

this flatness was not only because of anxiety and nervousness, it persisted even in students who appeared well-prepared. It reflected an absence of vocal training rather than an anxiety response. By mid-semester, following repeated drilling sessions and peer feedback, observations recorded a noticeable change by approximately 70%. Students started using their tone and speed more intentionally - slowing down at key points and using volume shifts to signal emphasis. Feedback from several students reported that the drilling exercises, which they had initially found artificial and uncomfortable, began to feel more natural when they understood the communicative function. Delivery scores on the vocal variety feature of the speech showed improvement from the first mini presentation to the final one across both class groups.

Pausing*:* Strategic pausing proved to be the feature that was the most visible in-class transformation. Early in the semester, students consistently treated silence as a problem to be avoided. Instructor reflections noted that when students forgot script or reached a transition point, the response was to fill the gap with 'um,' or repeated what they had just said. The drilling activities targeting pausing, in which students were asked to stop, hold silence for a full count of approximately 2-5 seconds. Teacher reflections from this period noted that normalizing silence and pausing required persistent reframing, both in explicit instruction and in the way the instructor modeled pausing during her own classroom lectures. By the latter half of the semester, however, pausing had become one of the most commented-upon improvements in peer feedback sessions. Students observing their classmates noted when a well-placed pause before a key point or transition gave them time to absorb what had just been said. Rubric scores on the pausing dimension reflected notable improvement, and student feedback at the end of the semester identified pausing as the delivery skill they felt most confident applying in their future speeches.

Intonation control: Intonation control showed equally meaningful improvement, though its development followed a somewhat different trajectory. Unlike vocal variety and pausing, in which students could see the relatively immediate results after practicing, intonation was a challenge for Vietnamese learners. Vietnamese is a tonal language in which pitch variation at the syllable level contains lexical meaning, and the relationship between pitch and meaning created interference patterns in students' English intonation production. My classroom observations from the drilling sessions on intonation revealed that students responded well to intonation practice, hearing and producing the same sentence with different intonation patterns and discussing the change in communicative meaning. Student feedback highlighted this approach as very useful, with several students noting that they had never considered intonation as a meaning-making resource in English, having focused on vocabulary and grammar accuracy. Teacher observations recorded an increase in the use of rising intonation to engage the audience, as well as more confident use of falling intonation to signal emphasis. Rubric scores showed improvement in intonation and stress across both class groups, a finding that points toward the need for continued attention to this feature in the public speaking course.

Student Engagement with AI-assisted contents

While delivery instruction was the most measurable outcome of the semester, the AI literacy phase produced a qualitatively important finding: across the two class groups, students tended to insert personal stories and specific examples from their experience into their speeches, adding them into their AI-generated scripts. This pattern was not seen from the beginning of the semester. In Mini presentation 1, approximately 70% of students maintained note-focused delivery with minimal eye contact. By the mid-term presentation, classroom observations recorded a significant increase in the frequency of personal stories / anecdotes

and locally specific examples in student speeches, with references to their own experiences and locally relevant contexts. These insertions showed that the students could reproduce the AI drafts and made them into their own versions. Teacher reflections from this period noted that this personal insertion was among the most communicatively effective. When a student paused to share a personal story or her own experience with the topic she was presenting, his/her audience would immediately engage, and the presenters could catch audience attention in the room easily. Student feedback at the end of the semester reinforced this observation. Students demonstrated greater resilience during Q&A sessions when they had personalized their material. They expressed confidence in their ability to add their insights to and take ownership of AI-generated materials, which represents a pedagogically productive relationship with these tools. The most compelling finding of the semester was the way in which the benefits of the two phases amplified each other in practice. The ownership of content seemed to motivate better delivery practice. The clearest illustration of this came in presentations where students had both revised and personalized their scripts and then applied their delivery training into their speeches. Teacher reflections noted that this pattern happened repeatedly in both class groups. Students who had improved delivery skills through drilling practices and feedback appeared to use those skills better with personalized contents. This interplay suggests that the dual-framework approach functions not merely as two separate instructional phases running in parallel, but as an integrated system in which each phase creates conditions for the other to operate more effectively.

5. Discussion & Recommendations

Discussion

Public speaking in EFL/ESL context, defined by Lucas (2009) is a strategic communication act, in which not only English proficiency but also critical thinking, creativity, and a logical structure for the speech are required. Public speaking demands that speakers present ideas clearly, systematically, and persuasively. Therefore, it is an effective approach to help EFL/ESL students develop communication awareness, build confidence, use English comprehensively, and, importantly, enhance their critical thinking skills (Zhang, 2009). The findings suggest productive directions for public speaking instruction in AI-integrated contexts. When AI tools can produce grammatically correct and coherent speech scripts, the focus on helping students build well-organized spoken texts remain important, but teachers need to give greater attention to personalization, delivery, and audience engagement. However, what AI cannot produce is an eloquent speaker, one who delivers personalized content with confidence and prosodic control. The dual-framework model implemented and proposed in this study and the findings suggest that it is a productive and practically viable foundation for rethinking oral communication pedagogy in EFL contexts.

The consistent improvement observed across all three delivery features - vocal variety, strategic pausing, and intonation control - supports a central argument of this paper that in the AI era, delivery instruction is a priority. When the script-production burden is done by AI assistance, class time and instructional energy can be redirected. The voice, the pause, and the body language are exactly what distinguishes a compelling speaker from a well-written script read aloud. The revision and personalization activities in Phase 1 did not ask students to stop using or refuse AI assistance but asked them to do something more demanding: to critically read the AI-generated draft, find out where it did not show their own knowledge and experience, and take steps to make it theirs. The result, as observed in classroom presentations and shown in student feedback, was what students invested in and took control of the content they are presenting, regardless of how that content was AI-drafted in the first place.

Recommendations

The findings and discussion above point toward a set of recommendations for TESOL practitioners teaching oral communication in AI-integrated EFL/ESL contexts. For public speaking courses, it is necessary for teachers to integrate AI literacy throughout speech preparation, make personalization a compulsory revision step, use peer feedback with clear observation form, and assess both content ownership and delivery quality. It is also useful if teachers clearly communicate to students from the first session that this course is about becoming an eloquent speaker, and that AI tools are welcome as a draft assistant.

The suggested activity sequence for a fifteen-week public speaking course is as follows. In the first 4 weeks of the course, it is recommended that the instructors introduce each delivery feature through focused drilling, with in-class practice presentations being used as drilling rather than assessed performances. From weeks 5 - 8, teachers can move to feedback sessions in which students observe and give feedback to each other's use of specific delivery features in practice presentations, using an observation form. From weeks 9-12, students need to identify one delivery strength and one priority for improvement in each presentation. From week 13-15, instead of commenting on individual features in isolation, students respond to the overall communicative effect of a presentation regarding the content and delivery.

Teachers need training and skills to integrate AI effectively into their practice. As AI becomes increasingly indispensable in education, understanding how to balance AI literacy with authentic oral communication is crucial. This will allow teachers to take advantage of AI's power and balance AI literacy with authentic oral communication.

6. Limitations

This study is restricted by its dependence on classroom observations, instructor's teaching reflections, and feedback from students; therefore, these data sources limit the generalizability of the findings beyond this specific institutional context and instructor. In addition, the findings are context-specific because the study involved two classes at one university and one instructor. The study adopted a teacher-researcher design, which may introduce subjectivity into both the instructional delivery and the observational data. The lack of detailed score reporting, limited reliability evidence for rubric scoring, limited information about the student feedback instrument, and the absence of a control group also mean that observed improvements in delivery and content ownership cannot be attributed merely to the dual-framework intervention, other factors, including students' self-practice and students' own development outside class may have contributed. Future studies should include a comparison group, multiple raters, pre- and post-performance scores, student in-depth interviews, and a clearer measure of AI literacy development and content ownership to strengthen the evidence for the model's effectiveness.

7. Conclusion

Vietnamese educational policy has signaled an ambitious shift in the role of English - from a foreign language taught as a subject to a medium of instruction, professional communication, and social participation embedded across the curriculum and the workplace (Ministry of Education and Training, 2017). This transition aims to produce a kind of speaker who can think in English, respond in English, use English to communicate naturally instead of rehearsed accuracy. The public speaking classroom, through the dual-framework model, is a foundation to build this kind of speaker - not because it only teaches English, but because it requires students to inhabit the language and use it to say something that matters to them in front of an engaged audience. The public speaking class, in the AI era, is for making good

speakers, not script producers or AI editors. Speakers who can stand confidently in front of an audience and deliver what they really want to say with the prosodic tools that make an audience engage in, stay with them, and believe that what is being said is worth hearing. The dual-framework model reported in this study provides one step toward that goal. This is not only a response to the use of AI in the classroom, it is also a response to the arrival of a new vision of what Vietnamese English education is for.

References

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Brazil, D. (1997). The communicative value of intonation in English. *Cambridge University Press.*

Derwing, T. M., & Munro, M. J. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. *John Benjamins.* https://doi.org/10.1075/lllt.42

Dong, Y. (2023). Revolutionizing academic English writing through AI-powered pedagogy: Practical exploration of teaching process and assessment. *Journal of Higher Education Research*, *4*(2), 52-67. https://doi.org/10.32629/jher.v4i2.1188

DuBose, J., & Marshall, D. (2023). AI in academic writing: Tool or invader. *Public Services Quarterly*, *19*(2), 125–130. https://doi.org/10.1080/15228959.2023.2185338

Ferris, D., & Tagg, T. (1996). Academic oral communication needs of EAP learners: What subject-matter instructors actually require. *TESOL Quarterly, 30*(1), 31–58. https://doi.org/10.2307/3587606

Fountoulakis, M. S. (2024). Evaluating the Impact of AI Tools on Language Proficiency and Intercultural Communication in Second Language Education. *International Journal of Second and Foreign Language Education*, *3*(1), 12–26. https://doi.org/10.33422/ijsfle.v3i1.768

Gilbert, J. B. (2008). Clear speech: Pronunciation and listening comprehension in North American English (4th ed.). *Cambridge University Press*.

Gilea, A, A., Katerina V.Melnova, & Gyuzelva K.Temirgaliyeva.(2025). The role of Artificial Intelligence in enhancing English speaking and listening skills in higher education. *Proceeding of International Conference on Social Science and Humanity*, 2(3), 870-879. https://doi.org/10.61796/icossh.v2i3.142

Joughin, G. (2010). The hidden curriculum revisited: A critical review of research into the influence of summative assessment on learning. *Assessment & Evaluation in Higher Education, 35*(3), 335–345. https://doi.org/10.1080/02602930903221493

King, J. (2002). Preparing EFL learners for oral presentations. *Dong Hwa Journal of Humanistic Studies, 4*, 401–418.

Levelt, W. J. M. (1989). Speaking: From intention to articulation. *MIT Press.*

Lucas, S. E., & Stob, P. (2020). The art of public speaking. *McGraw-Hill.*

Nguyen, T. T. T. (2025). Teacher scaffolding and collaborative learning in public speaking: Voices of students. *International Journal of English Language Studies*, *7*(1), 1–7. https://doi.org/10.32996/ijels.2025.7.1.1

Nguyen, Thi Anh Thu (2004). Prosodic transfer : the tonal constraints of Vietnamese acquisition of English stress & rhythm. PhD Thesis, School of English, Media Studies and Art History, The University of Queensland. https://doi.org/10.14264/107093

Nunan, D. (1989). Designing tasks for the communicative classroom. *Cambridge University Press.*

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W.,Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback on students' writing. *Learning and Instruction, 91*, Article 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Zhang, D. Y. (2009). Public speaking and the development of language competence. In L. Wang, & P. Li (Eds.), *English public speaking in global context: Challenges and innovations* (pp. 125–133). Foreign Language Teaching and Research Press.

Zhang, X., Ardasheva, Y., Egbert, J., & Ullrich-French, S. C. (2019). Building assessments for self-efficacy in English public speaking self-efficacy in China. *The Asia-Pacific Education Researcher*, *28*(5), 411-420. https://doi.org/10.1007/s40299-019-00441-9

#172 HUFLIT STUDENTS' PREPARATION FOR GRADUATION CERTIFICATES: PERCEIVED EFFECTIVENESS OF AI-GENERATED LEARNING SUPPORT

Quốc Bảo Mai[1,2], Thị Cẩm Vân Nguyễn[1,2], Đình Bảo Hân Nguyễn[1], & Nguyễn Nhựt Tuyết Hồ[1]

Abstract

Generative Artificial Intelligence is rapidly changing higher education. However, little is known about how students in Vietnamese private universities use these tools to prepare for important graduation certificates - such as TOEIC, VSTEP, and MOS - under strong academic pressure. This study addresses this gap by examining how Performance Expectancy, Effort Expectancy, and Academic Pressure influence students' AI Usage Behavior and their Perceived Effectiveness of learning. Using an explanatory sequential mixed-methods design, the study surveyed 260 English-majored juniors at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT) and interviewed six selected participants. Based on an extended Unified Theory of Acceptance and Use of Technology (UTAUT) framework, the quantitative results showed that Effort Expectancy (β = .470), Academic Pressure (β = .431), and Performance Expectancy (β = .307) all had positive and significant effects on AI Usage Behavior. AI Usage Behavior was also strongly associated with students' perceived learning effectiveness (β = .603). The interviews revealed a more nuanced picture. While students used GenAI as a learning support tool, pressure-driven reliance could produce a false sense of competence: students completed tasks with AI support but did not always build deep, independent understanding. Students also raised concerns about inaccurate AI-generated content and growing dependency. Overall, this study suggests that universities should move beyond outright bans and develop an AI-Integrated Academic Integrity Framework that promotes AI literacy and a human-in-the-loop approach, helping students use technology responsibly while still owning their learning.

***Keywords*:** Generative AI; AI Usage Behavior; Academic Pressure; Perceived Effectiveness; Graduation Certification; Extended UTAU

[1] *Undergraduate Student, Department of English Language, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT), Ho Chi Minh City, Vietnam*
[2] *Corresponding author*

1. Introduction

Higher education is changing as the Fourth Industrial Revolution automates routine cognitive tasks and raises demand for non-routine skills such as problem-solving and adaptability (Phan & Doan, 2019). This shift is closely tied to the emergence of Generative Artificial Intelligence (GenAI), particularly Large Language Models (LLMs) such as ChatGPT, which can produce human-like text responses (Bearman & Ajjawi, 2023; Kasneci et al., 2023). In Vietnam, this technological trend coincides with strict graduation requirements. Circular 12/2017/TT-BGDĐT requires higher education institutions to enforce foreign language proficiency (e.g., TOEIC, VSTEP) and computer skills (e.g., MOS) certifications as mandatory conditions for graduation (Ministry of Education and Training, 2017). Undergraduates at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT) therefore face a practical tension between using new AI tools and managing the pressure of these mandatory exit requirements.

GenAI is often described as a useful tool for self-regulated learning because it can provide quick and personalized feedback (Nguyen, N. T., 2025). However, a gap exists between what AI can support and how students actually use it. When students need to meet strict graduation requirements, they may use AI to reduce their cognitive effort rather than to deepen learning. This is particularly consequential in English Language Teaching (ELT), where language development typically requires sustained practice and active engagement. The result may be a competence gap: students produce answers that meet certificate requirements but do not fully understand or retain the underlying knowledge. Vo (2025) described this risk as a "false sense of competence" - students feel more capable than they actually are. Whether this type of AI use genuinely prepares students for work and further study, or only creates surface-level performance, remains unclear.

Problem Statement

The specific problem this study addresses is that, at HUFLIT, GenAI use and mandatory certificate requirements have grown together without clear evidence of how they interact. Because Circular 12/2017/TT-BGDĐT ties graduation directly to certificate scores, students have a strong incentive to use AI in whatever way gets them a passing result fastest. However, current UTAUT-based adoption research does not reveal whether that incentive pushes students toward AI as a genuine learning aid or as a shortcut around learning. This gap matters practically: if students pass TOEIC, VSTEP, or MOS examinations with heavy AI assistance but do not retain the underlying language or digital skills, the university risks certifying competencies that graduates do not actually hold, with consequences for their later academic and workplace performance.

A related issue is that existing technology-acceptance models such as UTAUT were not designed for high-stakes, deadline-driven certification settings, so they may understate how much Academic Pressure itself - separate from usefulness or ease of use - drives students toward AI-assisted shortcuts (Kasneci et al., 2023; Rudolph et al., 2023). Without empirical evidence on how Performance Expectancy, Effort Expectancy, and Academic Pressure jointly shape AI Usage Behavior, and on whether that usage produces genuine or only perceived learning gains, HUFLIT and similar institutions lack an evidence base for deciding between banning AI, ignoring it, or actively teaching students to use it well.

Research Objectives

To address this gap, this study examines how Performance Expectancy, Effort Expectancy, and Academic Pressure influence AI Usage Behavior and how AI Usage Behavior affects students' Perceived Effectiveness of learning at HUFLIT. Using an explanatory

sequential mixed-methods design, the study first tests the structural relationships among the main variables through a survey. It then uses interview data to explain students' AI use in greater depth, including their learning strategies, perceived test readiness, concerns about AI dependency, and ethical awareness.

Significance of the Study

This study contributes in several ways. Theoretically, it extends the UTAUT framework by adding Academic Pressure as a context-specific factor relevant to high-stakes education. Practically, it provides evidence for HUFLIT teachers and administrators about how students actually use AI for certification preparation. The study suggests that universities should move from outright AI bans toward an AI-Integrated Academic Integrity Framework. It also supports a shift to process-oriented assessment and encourages students to adopt a human-in-the-loop approach, in which AI serves as support while students retain responsibility for checking, evaluating, and developing their own understanding.

Scope and Delimitations

This study is scoped to junior undergraduate students in the Department of English Language at HUFLIT who are preparing for institutionally mandated graduation certificates (TOEIC, VSTEP, MOS, and second foreign language qualifications). Data were collected through a single cross-sectional survey (N = 260) and follow-up interviews (n = 6) within one semester, so the study captures students' perceptions and self-reported behavior at one point in time rather than change over time.

Several delimitations were set deliberately. First, the sample was limited to English-majored juniors because this cohort faces the most immediate exposure to certification deadlines; other majors, academic years, and institutions were excluded, which limits how far the findings can be generalized beyond similar Vietnamese private-university contexts. Second, Perceived Effectiveness is defined as students' self-reported sense of learning support and test readiness, not as independently verified test scores or objective learning gains. Third, the extended UTAUT model includes three antecedents - Performance Expectancy, Effort Expectancy, and Academic Pressure - chosen for their relevance to high-stakes certification contexts; other UTAUT constructs (e.g., Social Influence, Facilitating Conditions) and individual differences (e.g., prior AI literacy, baseline English proficiency) fall outside the scope of this study. Finally, the qualitative phase was limited to six purposively selected interviewees, which supports interpretive depth but not statistical generalization of the qualitative themes.

2. Literature Review

Overview of Generative Artificial Intelligence in Education

Generative Artificial Intelligence refers to AI systems that create new content-text, images, code - by learning patterns from large datasets (Bearman & Ajjawi, 2023). Within higher education, LLMs such as ChatGPT and Gemini have moved quickly from novelty to widely used tools, offering benefits such as personalized feedback and support for self-regulated learning while simultaneously raising concerns about academic integrity and independent thinking (Kasneci et al., 2023; Chan & Tsi, 2023). For English as a Second Language (ESL) learners, GenAI can act as a responsive tutor that provides immediate corrective feedback and explains grammar or vocabulary. This accessible, individualized support is especially useful for non-native speakers preparing for standardized proficiency assessments (Nguyen, N. T., 2025; Almogren et al., 2024).

Theoretical Frameworks

The Evolutionary Trajectory: From TAM to UTAUT

Technology adoption research is grounded in the Technology Acceptance Model (TAM), proposed by Davis (1989), which explained behavioral intentions through Perceived Usefulness and Perceived Ease of Use. As digital environments grew more complex, Venkatesh et al. (2003) synthesized eight earlier models to create the Unified Theory of Acceptance and Use of Technology (UTAUT), which improved predictive power by identifying Performance Expectancy (PE) and Effort Expectancy (EE) as key determinants of technology use. Recent studies in the GenAI domain confirm these relationships: Parveen et al. (2024) used Structural Equation Modeling to show that PE drives ChatGPT adoption, and Strzelecki (2024) found Performance Expectancy to be the strongest predictor of students' behavioral intention to use ChatGPT across a large sample of university students.

Academic Pressure in High-Stakes Certification Contexts

Academic Pressure refers to the stress students experience when they must meet academic deadlines, pass important exams, and satisfy graduation requirements. In this study, Academic Pressure is specifically tied to certificate requirements such as TOEIC, VSTEP, MOS, and second foreign language qualifications. This pressure differs from general academic stress because it is linked to graduation conditions, fixed timelines, exam fees, retake costs, and the risk of delayed graduation. When facing these high-stakes requirements, students may use GenAI not only because it is useful or easy to use, but because it helps them save time, reduce stress, and prepare more quickly. Nguyen, A. H. (2025) showed that academic pressure is associated with students' AI use and concerns about academic integrity. Similarly, Beyersdorf et al. (2025) found that students often consider using GenAI when they face time pressure and difficult academic tasks. Academic Pressure is therefore included in this study as a context-specific factor that may strongly influence AI Usage Behavior among HUFLIT students.

Synthesis: The Extended Research Model

Although UTAUT provides a useful structure for explaining technology acceptance, its generic application may not fully capture the contextual pressures of Vietnamese higher education. Drawing on Nguyen, A. H. (2025), this study extends the UTAUT framework by incorporating Academic Pressure as a context-specific antecedent. The proposed model posits that AI adoption in this setting is shaped not only by students' perceived usefulness and ease of use, but also by institutional pressure surrounding high-stakes graduation requirements. Accordingly, Performance Expectancy, Effort Expectancy, and Academic Pressure are examined in relation to AI Usage Behavior, which is then linked to Perceived Effectiveness. Direct paths from these antecedents to Perceived Effectiveness are also estimated to assess whether AI Usage Behavior functions as a partial mediator.

Review of Related Studies

The following subsections review empirical studies that support the logic of the research model. Research on GenAI capabilities in examinations establishes the tool context. Studies on effectiveness versus dependency explain why Perceived Effectiveness must be interpreted carefully. Research on hallucination and academic integrity frames the qualitative themes.

AI Capabilities in High-Stakes Examinations

The strong performance of GenAI on examination tasks has moved from theoretical concern to empirical reality, raising questions about the reliability of traditional assessments. Wu et al. (2025) found that GPT-4o performed significantly better than human candidates on Taiwanese medical licensing exams. Ali et al. (2023) showed that AI tools achieved high proficiency on neurosurgery oral board questions. In the Vietnamese context, Pham (2024) found that Google Bard outperformed the average human score on the National High School Geography exam. These studies suggest that GenAI can support task completion and test practice in ways that may not reflect students' independent understanding - a concern directly relevant to institutional exit qualifications such as TOEIC and VSTEP.

Effectiveness versus Dependency in Language Learning

The use of LLMs in language learning creates both opportunities for perceived learning effectiveness and risks of cognitive dependency. While Nguyen, N. T. (2025) reported that AI can support vocabulary learning and structural comprehension, students may also over-rely on AI for tasks requiring independent thinking. Using latent profile analysis, Stojanov et al. (2024) identified a subgroup of students whose heavy reliance on ChatGPT was associated with weaker independent study habits, lending empirical support to the dependency concern. Bearman and Ajjawi (2023) similarly argued that students need to learn how to work critically with AI rather than use it as a replacement for their own judgment. In this study, Perceived Effectiveness is therefore defined as students' self-reported sense of learning support and test readiness, not as direct evidence of objective learning achievement.

Challenges and Academic Integrity

GenAI also raises ethical and information-reliability concerns. Anh-Hoang et al. (2025) identified "algorithmic hallucinations"-the phenomenon where AI confidently generates plausible but factually incorrect content-as a key challenge requiring users to verify AI-generated information carefully. Perkins (2023) argued that institutions cannot rely on plagiarism-detection tools alone and need updated academic integrity policies that explicitly address AI-assisted work. For this study's theoretical structure, Nguyen, A. H. (2025) established a correlation between Academic Pressure and the propensity for academic integrity transgressions among undergraduates. This suggests a survival-oriented pattern of AI use, where students facing graduation deadlines may prioritize certification outcomes over the learning process.

Hypothesis Development

Synthesizing the literature with the extended UTAUT framework, this study advances the following hypotheses.

Performance Expectancy (PE) and AI Usage Behavior (H1)

Performance Expectancy reflects the extent to which students believe that using a system will improve their academic performance. Originating from the Perceived Usefulness construct in Davis's (1989) TAM, PE was identified by Venkatesh et al. (2003) as the strongest predictor of behavioral intention in UTAUT. In the context of GenAI for certification preparation, students are motivated by practical benefits. Parveen et al. (2024) demonstrated through SEM that PE drives ChatGPT adoption, with students viewing the tool as a productivity enhancer. Nguyen, N. T. (2025) found a strong positive relationship between AI content generation capabilities and perceived learning effectiveness among HUFLIT students. Dang et al. (2024) reported that ChatGPT has become a common learning and research support

tool among students in Ho Chi Minh City. Strzelecki (2024) confirmed PE as the strongest predictor of students' intention to use ChatGPT in university settings. If HUFLIT students perceive AI as a tool that can simulate TOEIC mock exams or clarify grammar concepts more effectively than traditional resources, this expectation should directly increase their usage.

H1: Performance Expectancy has a positive and significant influence on AI Usage Behavior.

Effort Expectancy (EE) and AI Usage Behavior (H2)

Effort Expectancy refers to the perceived ease of use associated with a technology. Rooted in Davis's (1989) Perceived Ease of Use, this construct suggests that intuitive systems encourage frequent use, whereas complex technologies encounter user resistance (Venkatesh et al., 2003). GenAI tools are distinctive because their natural language interfaces substantially lower technical barriers, especially for non-IT majors. Parveen et al. (2024) noted that ChatGPT's conversational design significantly reduces students' cognitive load. Nguyen, N. T. (2025) found that effortless content generation encourages sustained engagement. Bearman and Ajjawi (2023) characterized AI as a user-friendly "black box" where simple prompts yield complex outputs, thereby increasing students' willingness to integrate it into their study routines. If HUFLIT students find interacting with AI to be straightforward and effortless, their usage frequency should increase.

H2: Effort Expectancy has a positive and significant influence on AI Usage Behavior.

Academic Pressure (AP) and AI Usage Behavior (H3)

Moving beyond traditional models that focus primarily on technological features, this study incorporates Academic Pressure as a critical determinant. Defined as the psychological strain from strict deadlines, elevated expectations, and the fear of academic failure, this construct draws on Agnew's (1992) General Strain Theory, which holds that individuals adopt specific coping strategies under stress. In Vietnamese higher education, such pressure can act as a catalyst for AI adoption. Nguyen, A. H. (2025) identified a direct correlation between academic pressure and AI usage, suggesting that students confronted with impending VSTEP deadlines often prioritize efficiency over independent study. Beyersdorf et al. (2025) noted that undergraduates frequently justify GenAI use when subjected to time constraints and high-stakes evaluations. Under these conditions, the threat of delayed graduation may transform AI from a learning aid into a coping mechanism.

H3: Academic Pressure has a positive and significant influence on AI Usage Behavior.

AI Usage Behavior (UB) and Perceived Effectiveness (PEA) (H4)

Perceived Effectiveness relates to students' self-assessment of learning support, perceived skill improvement, and confidence regarding test readiness. Usage Behavior functions as the critical bridge: as Reich (2020) emphasized, access to technology alone does not guarantee learning without active engagement. Empirical evidence connects active AI integration with higher perceived competence. In high-pressure academic environments, sustained interaction with AI tools shapes how students evaluate their own proficiency (Nguyen, A. H., 2025; Beyersdorf et al., 2025). As students use GenAI to navigate academic demands (Klitgaard, 2025), their sustained engagement is associated with their subjective perception of academic success.

H4: AI Usage Behavior has a positive and significant influence on Perceived Effectiveness.

In addition to the hypothesized paths, direct paths from Performance Expectancy, Effort Expectancy, and Academic Pressure to Perceived Effectiveness were estimated. These paths were included to determine whether AI Usage Behavior served as a full or partial mediator in the structural model.

The Mediating Role of AI Usage Behavior (H5, H6, H7)

Central to this research framework is the mediating role of AI Usage Behavior. Drawing on mediation analysis principles (Baron & Kenny, 1986), an independent variable may influence a dependent variable through an intervening mechanism. In this study, AI Usage Behavior acts as a behavioral pathway through which students' beliefs about usefulness, perceived ease of use, and academic pressure are translated into perceived learning outcomes. Because the structural model also estimates direct paths from the three antecedents to Perceived Effectiveness, the analysis can determine whether the mediation is full or partial. In a high-stakes certification context, pressure may directly shape students' confidence and perceived readiness, while also encouraging more frequent AI-assisted study.

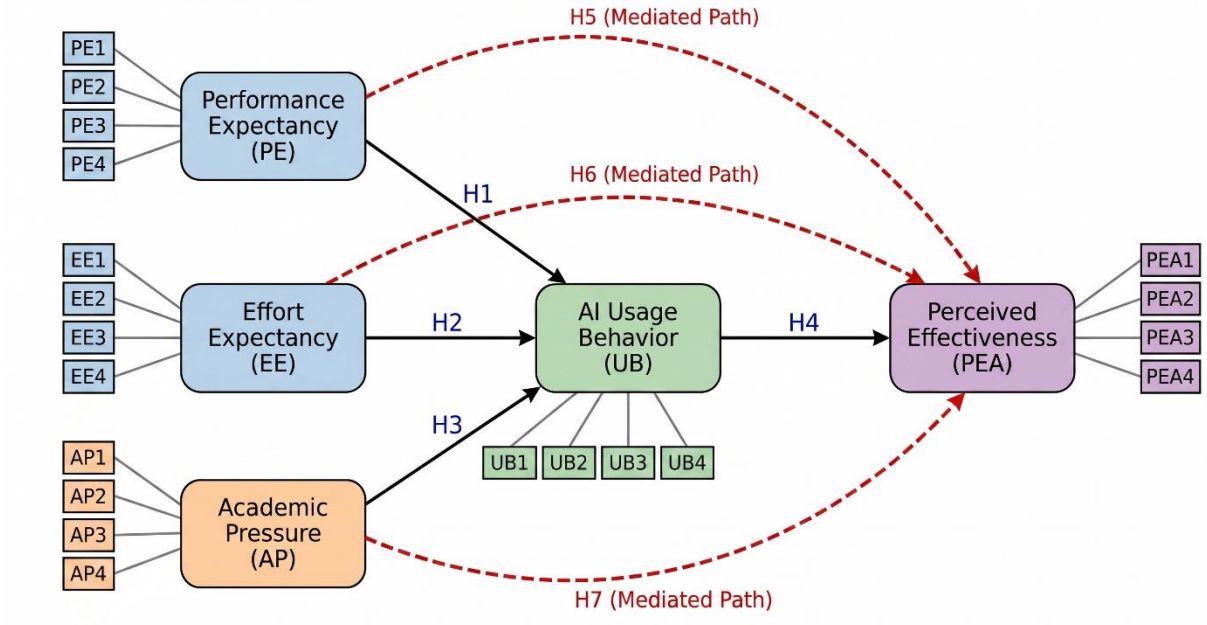
H5: AI Usage Behavior mediates the relationship between Performance Expectancy and Perceived Effectiveness.

H6: AI Usage Behavior mediates the relationship between Effort Expectancy and Perceived Effectiveness.

H7: AI Usage Behavior mediates the relationship between Academic Pressure and Perceived Effectiveness.

Figure 6

Proposed Extended UTAUT Research Model



3. Methodology

Research Design

This research employs an explanatory sequential mixed-methods research design (Creswell, 2014). The quantitative phase tests the hypothesized structural relationships within the extended UTAUT model. The qualitative phase then follows to explore participants' lived

experiences, particularly their sense of competence, ethical concerns, and possible dependency on AI. This two-phase approach allows the interview findings to explain, contextualize, and nuance the survey results.

Participants and Research Context

The target population consists of junior undergraduate students enrolled in the Department of English Language at HUFLIT, Vietnam. Using purposive sampling, 260 participants were recruited for the quantitative phase. This cohort was selected because juniors face the most immediate exposure to graduation certification deadlines (e.g., TOEIC, VSTEP, MOS) and the associated academic pressure. The sample size of 260 exceeds the minimum requirement calculated using the "10-times rule" commonly applied in PLS-SEM (Hair et al., 2019).

For the qualitative phase, six participants were purposively selected from the survey respondents. Selection criteria focused on maximizing variation in self-reported GenAI use patterns (from highly dependent users to strategic users) and willingness to discuss their learning experiences in depth. All participants provided informed consent, participation was voluntary, and data anonymity was guaranteed throughout the study.

Instrumentation and Data Collection

The quantitative data were collected using a closed-ended questionnaire based on a five-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree). Core constructs were measured by adapting pre-validated scales to the context of LLM integration. Items for Performance Expectancy and Effort Expectancy were adapted from Venkatesh et al. (2003). Items for Academic Pressure, AI Usage Behavior, and Perceived Effectiveness were synthesized based on Nguyen, A. H. (2025). Before formal deployment, the instrument was pilot-tested with 30 undergraduates to evaluate semantic clarity and internal consistency.

Formal data collection took place over a four-week period using an online survey distributed via Google Forms. Responses with incomplete data or straight-lining patterns were excluded to ensure data quality. Qualitative data were collected through semi-structured interviews lasting approximately 30–45 minutes each, conducted in Vietnamese to allow participants to express their thoughts accurately. All interviews were audio-recorded with explicit permission. The interview protocol was designed to elicit narratives about AI use for specific tasks, self-reflection on technological dependency, and views on the role of lecturers.

Analytical Procedures

Quantitative data underwent screening before analysis using IBM SPSS Statistics and SmartPLS 4.0. Descriptive statistics were generated to outline the demographic profile and baseline AI usage patterns. Partial Least Squares Structural Equation Modeling (PLS-SEM) was then conducted using SmartPLS 4.0. PLS-SEM was selected over Covariance-Based SEM (CB-SEM) for three reasons: its robustness with complex predictive models, its suitability for exploratory frameworks (such as extending UTAUT with Academic Pressure), and its relaxed distributional assumptions (Hair et al., 2019). Common Method Bias (CMB) was assessed using full collinearity Variance Inflation Factors (VIF).

The analysis proceeded in two stages: first, evaluating the measurement model's convergent and discriminant validity; second, testing the structural model's direct and indirect paths. A non-parametric bootstrapping procedure with 5,000 subsamples (two-tailed, 95% bias-corrected confidence intervals) was used to evaluate significance. Direct paths from the three antecedents to Perceived Effectiveness were included to assess mediation type.

For the qualitative phase, Vietnamese interview recordings were transcribed verbatim and translated into English. Transcripts were analyzed using MAXQDA professional software following an inductive thematic analysis approach. Two independent researchers coded the data; initial codes were generated independently, compared, and refined through discussion until inter-coder consensus was reached. Qualitative themes were then integrated with the quantitative results to provide a comprehensive explanation of students' AI adoption behaviors.

4. Results

Demographic Profile of the Respondents

Table 1 outlines the demographic and academic profile of the 260 participants. Females constituted 57.7% of the sample; males accounted for 42.3%. The majority of participants were preparing for multiple certifications simultaneously, with TOEIC (90.8%) and Second Foreign Languages (78.5%) showing the highest prevalence, followed by VSTEP (73.5%). Regarding baseline proficiency, the distribution suggests a sample with moderate to good foundational competence.

Table 41

Demographic and Academic Profile of Respondents

Demographic Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	110	42.3%
	Female	150	57.7%
Target Certification *(Multiple selection permitted)*	English: TOEIC (2 or 4 skills)	236	90.8%
	English: IELTS	164	63.1%
	English: VSTEP	191	73.5%
	Informatics: MOS (Excel/Word/PPT)	125	48.1%
	Second Foreign Language	204	78.5%
Current Proficiency	Very Low	19	7.3%
	Low	20	7.7%
	Average	61	23.5%
	Good	95	36.5%
	Excellent	65	25.0%

Note: Multiple selections were permitted for target certifications.

Usage Habits and Metrics (UHM)

To examine how students use GenAI in their independent study, usage habits and engagement patterns were tabulated. GenAI tools are widely and frequently used among this cohort. ChatGPT is the most popular platform (87.3%), followed closely by Grammarly/Quillbot (86.5%) and Gemini (81.2%). A large share of students reported spending considerable time on AI-assisted academic tasks daily.

Table 42

Primary AI Tools Utilized by Respondents

Tool	Frequency (n)	Percentage (%)
ChatGPT (OpenAI)	227	87.3%
Grammarly / Quillbot	225	86.5%
Gemini (Google)	211	81.2%
Copilot (Microsoft)	162	62.3%
Elsa / Duolingo	162	62.3%

*Multiple selections were permitted.

Descriptive Statistics of the Measurement Items

Descriptive data were examined for central tendency and dispersion. Mean scores ranged from 3.231 to 3.773, indicating moderately to highly positive attitudes toward AI use. All skewness and kurtosis values fell within ±1.0, confirming acceptable univariate normality for PLS-SEM analysis. Full item-level statistics are available upon request.

Table 43

Descriptive Statistics of Measurement Items

Construct	Item	Mean	Std. Deviation	Skewness	Excess Kurtosis
PE	PE1 - PE4	3.438 - 3.473	0.775 - 0.837	-0.204 to 0.210	-0.501 to -0.055
EE	EE1 - EE4	3.658 - 3.704	0.800 - 0.818	-0.288 to -0.072	-0.251 to 0.055
AP	AP1 - AP4	3.488 - 3.546	0.776 - 0.860	-0.283 to -0.206	-0.314 to 0.009
UB	UB1 - UB4	3.723 - 3.773	0.718 - 0.734	-0.279 to -0.126	-0.223 to 0.252
PEA	PEA1 - PEA4	3.231 - 3.269	0.723 - 0.772	-0.088 to 0.148	-0.368 to 0.066

Measurement Model: Reliability and Validity

The measurement model was evaluated for indicator reliability, internal consistency, convergent validity, and discriminant validity. As shown in Table 4, all constructs demonstrated strong reliability, with Composite Reliability (CR) values exceeding the 0.70 threshold. Convergent validity was confirmed through Average Variance Extracted (AVE) values above the 0.50 benchmark.

Table 44

Measurement Model: Indicator Reliability, Internal Consistency, and Convergent Validity

Construct	Items	Outer Loadings	Cronbach's Alpha	Composite Reliability (CR)	Average Variance Extracted (AVE)
Academic Pressure (AP)	AP1	0.872	0.894	0.926	0.759
	AP2	0.873			
	AP3	0.882			
	AP4	0.858			
Effort Expectancy (EE)	EE1	0.883	0.905	0.934	0.779
	EE2	0.864			
	EE3	0.891			
	EE4	0.892			
Performance Expectancy (PE)	PE1	0.883	0.904	0.933	0.776
	PE2	0.886			
	PE3	0.886			
	PE4	0.869			
Perceived Effectiveness (PEA)	PEA1	0.839	0.869	0.911	0.718
	PEA2	0.849			
	PEA3	0.852			
	PEA4	0.850			
Use Behavior (UB)	UB1	0.824	0.868	0.910	0.717
	UB2	0.874			
	UB3	0.865			
	UB4	0.823			

Discriminant validity was assessed using the Heterotrait-Monotrait (HTMT) ratio of correlations (Henseler et al., 2015). As shown in Table 5, all HTMT values were below the 0.90 threshold (maximum value: 0.871 between UB and PEA), confirming that each construct captures a distinct phenomenon.

Table 45

Discriminant Validity (HTMT Ratio)

Construct	AP	EE	PE	PEA	UB
AP					
EE	0.179				
PE	0.029	0.057			
PEA	0.499	0.528	0.345		
UB	0.576	0.627	0.371	0.871	

Structural Model and Hypotheses Testing

A non-parametric bootstrapping approach with 5,000 subsamples was executed in SmartPLS 4.0. As shown in Table 6, Performance Expectancy (β = .307, p < .001), Effort Expectancy (β = .470, p < .001), and Academic Pressure (β = .431, p < .001) all had significant positive effects on AI Usage Behavior, supporting Hypotheses 1, 2, and 3. The path from Usage Behavior to Perceived Effectiveness was also strong and significant (β = .603, p < .001), supporting Hypothesis 4.

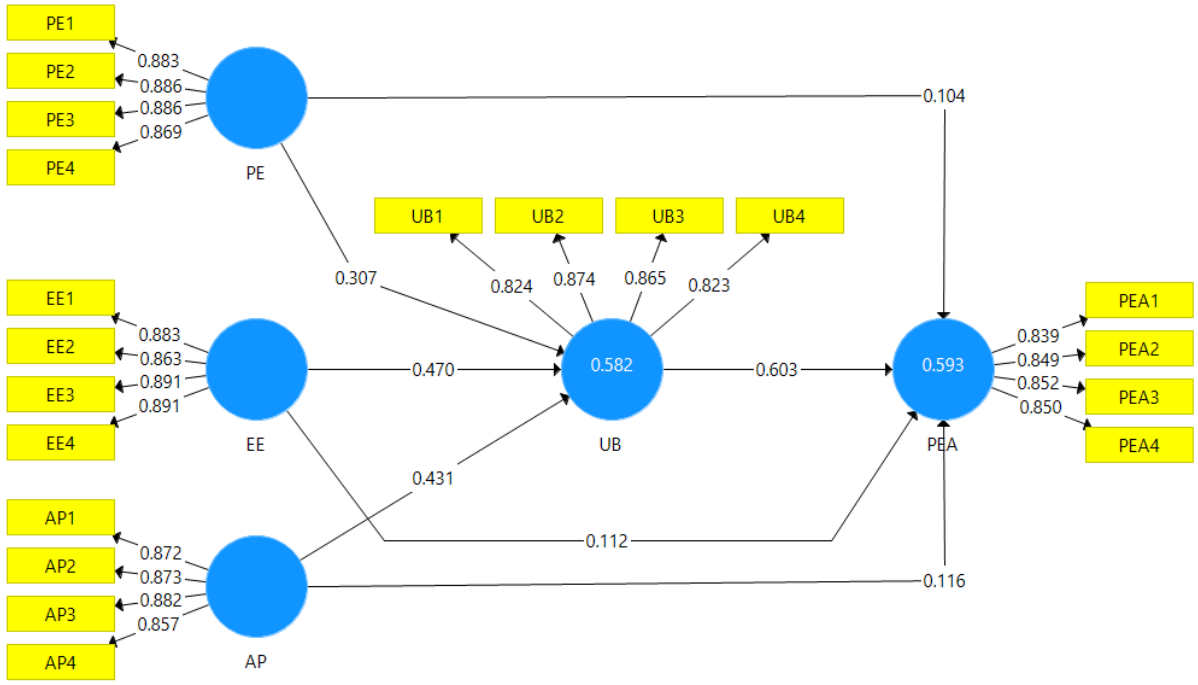
Table 46

Structural Model Assessment (Direct Effects)

Hypothesis	Structural Path	*β*	*t*-Statistic	*p*-Value	95% CI (BC)	Decision
H1	PE -> UB	.307	7.822	< .001	[.225, .381]	Supported
H2	EE -> UB	.470	13.265	< .001	[.397, .536]	Supported
H3	AP -> UB	.431	11.148	< .001	[.350, .504]	Supported
H4	UB -> PEA	.603	9.754	< .001	[.480, .722]	Supported
-	PE -> PEA	.104	2.284	< 0.05	[.012, .191]	Supported
-	EE -> PEA	.112	2.191	< 0.05	[.009, .211]	Supported
-	AP -> PEA	.116	2.256	< 0.05	[.014, .215]	Supported

Figure 7

PLS-SEM Structural Model and Hypotheses Testing



The specific indirect effects were then examined to clarify the mediating role of AI Usage Behavior (Table 7). AI Usage Behavior significantly mediated the relationship between Performance Expectancy and Perceived Effectiveness (H5: $\beta = .185$, $t = 6.267$, $p < .001$). The mediating pathways for Effort Expectancy (H6: $\beta = .283$, $t = 7.419$, $p < .001$) and Academic Pressure (H7: $\beta = .260$, $t = 7.516$, $p < .001$) were also significant. Because the corresponding direct paths to Perceived Effectiveness were also significant, the results indicate partial rather than full mediation.

Table 47

Structural Model Assessment (Direct Effects and Confidence Partial mediation)

Hypothesis	Mediating Pathway	β	*t*-Statistic	*p*-Value	95% CI (BC)	Decision
H5	PE -> UB -> PEA	.185	6.267	< .001	[.131, .248]	Patrial
H6	EE -> UB -> PEA	.284	7.419	< .001	[.211, .362]	Patrial
H7	AP -> UB -> PEA	.260	7.516	< .001	[.197, .334]	Patrial

Explanatory Power (R^2) and Effect Size (f^2)

The R^2 values for the two endogenous constructs-AI Usage Behavior and Perceived Effectiveness - were .582 and .593, respectively. The exogenous variables explain 58.2% of the variance in Usage Behavior and 59.3% of the variance in Perceived Effectiveness. Both values exceed the 0.50 threshold, indicating moderate to strong explanatory power (Hair et al., 2019).

Effect size (f^2) analysis revealed that Effort Expectancy ($f^2 = .514$) and Academic Pressure ($f^2 = .433$) exert large effects on AI Usage Behavior. Performance Expectancy demonstrated a moderate effect ($f^2 = .225$). AI Usage Behavior exhibited a large effect on Perceived Effectiveness ($f^2 = .373$). Together, these findings confirm that AI Usage Behavior serves as a significant mediating mechanism through which PE, EE, and AP are translated into Perceived Effectiveness.

Model Fit

The overall goodness-of-fit was evaluated using the Standardized Root Mean Square Residual (SRMR). The estimated SRMR value is 0.046, which is below the conservative threshold of 0.08 (Hu & Bentler, 1999), confirming an excellent fit between the theoretical model and the empirical data.

Out-of-Sample Predictive Power (PLSpredict)

The PLSpredict procedure with 10 folds and 10 repetitions was used to assess out-of-sample predictive capability (Shmueli et al., 2019). All Q^2 values for the endogenous constructs were strictly greater than zero (ranging from 0.267 to 0.421), confirming that the PLS path model outperforms the naïve benchmark. Furthermore, the Root Mean Square Error (RMSE) values from the PLS-SEM model were consistently lower than those from the naïve linear regression model (LM) across all indicators. This confirms that the proposed structural model has high predictive power.

Qualitative Research Findings: In-depth Interview Analysis

Semi-structured interviews with six HUFLIT students were conducted to provide interpretive depth to the quantitative findings. The thematic analysis aimed to clarify the psychological motivations, usage strategies, perceived learning outcomes, and ethical concerns associated with GenAI use for certification preparation. Five major themes emerged: (1) academic pressure as a coercive driver, (2) the distinction between learning scaffolding and cognitive offloading, (3) a possible false sense of competence, (4) information control and ethical boundaries, and (5) the continuing importance of faculty guidance.

Theme 1: Academic Pressure as a Coercive Driver

The interview findings reinforced the quantitative result for Hypothesis 3. Participants consistently described graduation requirements as a source of urgency that pushed them toward AI-supported study. Certifications such as TOEIC, VSTEP, and MOS were perceived not merely as academic milestones but as institutional barriers affecting graduation timelines. Interviewee 1 stated that because graduation was approaching, they would naturally use any AI tool that could help them graduate on time. This response suggests that AI adoption is often motivated by the perceived need to cope with unavoidable academic demands rather than by curiosity alone.

At the same time, students did not necessarily view AI as a complete replacement for learning. Interviewee 4 explained that AI helped reduce pressure to some extent but emphasized that it should remain a support tool. However, under severe time constraints, AI use shifted from learning scaffold to task-completion shortcut. Interviewee 1's admission of copy-pasting AI outputs for class assignments illustrates how systemic pressure may encourage efficiency-oriented behaviors that do not contribute to sustainable learning. The qualitative evidence thus suggests that academic pressure operates as a coercive force, transforming AI from an optional study aid into a survival mechanism.

Theme 2: Learning Scaffolding vs. Cognitive Offloading

The interviews revealed a clear behavioral distinction between students who used AI actively as a learning scaffold and those who relied on it passively for cognitive offloading. In the first pattern, AI served as a tool for organizing, practicing, and deepening learning. Interviewee 2 used AI to generate customized practice exercises (multiple-choice and Yes/No questions) for independent study outside the classroom. Interviewee 4 used AI to request conceptual outlines rather than complete essays. These practices show that when students control the learning process and use AI selectively, technology can support autonomy and structured preparation.

However, some participants expressed concern that excessive reliance on AI could weaken independent reasoning. Interviewee 5 stated that using AI too much could lead to dependency and negatively affect critical thinking. This concern reflects cognitive offloading-transferring mental tasks such as idea generation, grammar correction, and translation to AI systems. Although AI may increase speed and convenience, it can also reduce opportunities for deep engagement with learning material. The educational value of AI therefore depends less on whether students use it and more on how they use it.

Theme 3: A Possible False Sense of Competence

One of the most significant qualitative findings was a possible false sense of competence. Participants frequently described a gap between feeling capable with AI support and feeling uncertain when required to perform independently. This theme helps explain the

quantitative finding that AI Usage Behavior strongly predicted Perceived Effectiveness. While students may perceive AI-supported learning as effective, the interviews suggest that this effectiveness can be temporary and technologically mediated rather than evidence of independent mastery.

Interviewee 3 warned that AI may create the feeling that one's ability is higher than it actually is. When students allow AI to solve problems or generate answers, they may not internalize the knowledge needed for real exam conditions without AI. Interviewee 2 acknowledged that forgetting answers after AI-supported practice is natural, because AI only responds to the user's input. This highlights the temporary nature of AI-mediated performance: students may complete tasks successfully with AI assistance yet struggle to retain or reproduce the knowledge independently.

Theme 4: Information Control and Ethical Boundaries

Despite frequent AI use, participants did not demonstrate blind trust in AI-generated responses. Many showed algorithmic skepticism, particularly regarding inaccurate or fabricated information. Interviewee 6 emphasized that AI-generated information could not always be considered reliable and needed to be verified through textbooks or Google searches.

Students also articulated their own ethical boundaries. Interviewee 4 distinguished between asking AI to suggest ideas (acceptable) and having AI write an entire essay (unacceptable). Interviewee 3 raised concerns about inequality, noting a possible skill gap between students who use AI effectively and those who do not. This observation suggests that AI literacy is becoming an important academic resource, and unequal access to effective usage strategies may create new disparities. Overall, students were concerned not only with convenience but also with reliability, fairness, and appropriate academic conduct.

Theme 5: The Enduring Role of Faculty

Although students acknowledged AI's usefulness, they placed stronger trust in human teachers. Interviewee 5 stated that students trust faculty the most because teachers have practical experience. Interviewee 1 similarly prioritized lecturers' pedagogical expertise over algorithmically generated explanations. These responses suggest that AI has not displaced the authority of human instruction; students view AI as a supplementary tool whose value must be guided and evaluated by educators.

Participants also expressed a need for institutional support and formal AI literacy training. Interviewee 2 suggested that HUFLIT should organize targeted training sessions to help students use AI more appropriately. Such training could help students move from passive consumption of AI-generated answers toward more critical and reflective usage. This finding carries important pedagogical implications: rather than prohibiting AI outright, universities should provide structured guidance on ethical, effective, and learning-oriented AI practices. Faculty remain central not only as sources of academic knowledge but as mediators who help students develop the evaluative judgment needed to use AI responsibly.

5. Discussion

The findings of this study provide a contextualized understanding of how HUFLIT undergraduates adopt GenAI when preparing for graduation-related certifications. Overall, the results validate the extended UTAUT model: students' AI usage is driven by both traditional technological factors - Effort Expectancy and Performance Expectancy - and by Academic Pressure as a contextual factor. While Effort Expectancy ($\beta = .470$) exerted the strongest influence, followed by Academic Pressure ($\beta = .431$) and Performance Expectancy ($\beta = .307$),

the emergence of Academic Pressure as a key antecedent is particularly meaningful. HUFLIT students must obtain certifications such as TOEIC, VSTEP, MOS, and second foreign language qualifications before graduation. In such a context, AI is perceived not merely as an optional innovation but as a practical response to institutional pressure, time constraints, and the fear of delayed graduation.

The support for Hypotheses 1 and 2 confirms that Performance Expectancy and Effort Expectancy significantly influence AI Usage Behavior, consistent with the original UTAUT framework (Venkatesh et al., 2003). The strong effect of Effort Expectancy highlights that the conversational, frictionless interfaces of GenAI tools are a major adoption motivator. This aligns with Parveen et al. (2024), who reported PE as a key driver of ChatGPT adoption, and with Strzelecki (2024), who identified PE and EE as the strongest predictors of ChatGPT use among university students.

The significant effect of Academic Pressure on AI Usage Behavior (Hypothesis 3) confirms the value of extending UTAUT with a context-specific antecedent. While the original UTAUT emphasizes technological acceptance, this study shows that in high-stakes academic settings, technology adoption is also reinforced by coercive educational conditions. This aligns with Agnew's (1992) General Strain Theory, which suggests that individuals under pressure adopt coping strategies to manage stress. In the HUFLIT context, graduation certificates function as institutional gatekeeping mechanisms, and students may turn to AI not only for learning efficiency but also to reduce anxiety and avoid delayed graduation. This finding is consistent with Nguyen, A. H. (2025), who identified academic pressure as a factor associated with AI use and academic integrity concerns. The qualitative evidence further explains this relationship: students' AI usage is embedded in a survival-oriented academic logic, where the immediate goal is often to pass requirements rather than to develop stable, long-term competence.

AI Usage Behavior had a strong positive effect on Perceived Effectiveness (Hypothesis 4). Students who used AI more frequently reported higher confidence and stronger perceptions of learning effectiveness. At first glance, this supports the perceived value of AI-assisted learning. However, when integrated with the interview data, a more complex picture emerges. The qualitative evidence suggests that perceived effectiveness may sometimes reflect a false sense of competence: participants reported that AI made them feel capable during practice, but they also recognized that this confidence might not hold in real exam conditions without AI. This aligns with Bearman and Ajjawi (2023), who warned that excessive reliance on AI might foster superficial fluency without deep cognitive assimilation.

Interestingly, AI Usage Behavior showed higher mean scores than Perceived Effectiveness. This suggests that frequent AI use does not automatically translate into strong perceived learning gains. This pattern is consistent with the interview findings, where some students used AI for speed and task completion while remaining uncertain about their independent mastery.

The supported mediation effects (Hypotheses 5, 6, and 7) further clarify these mechanisms. AI Usage Behavior significantly mediated the relationships between all three antecedents and Perceived Effectiveness. For Academic Pressure specifically, this means that pressure may shape students' Perceived Effectiveness both directly and indirectly. Indirectly, it encourages more frequent AI use, which then strengthens students' perceived effectiveness. This finding is central to the study's theoretical contribution: in high-pressure educational contexts, perceived learning effectiveness may be constructed through repeated AI-assisted performance rather than through independently verified mastery. The relationship between AI

use and learning effectiveness should therefore be interpreted with caution, particularly when effectiveness is measured through subjective self-perception.

The qualitative themes also reveal a clear distinction between productive and problematic forms of AI use. Some students used AI as a learning scaffold-generating mock tests, requesting outlines, clarifying grammar points, or practicing independently. These practices are consistent with the view that GenAI can support autonomous learning and personalized feedback (Nguyen, N. T., 2025; Almogren et al., 2024). However, other students expressed concern that excessive AI use could weaken critical thinking and create dependency. This distinction echoes Bearman and Ajjawi's (2023) argument that higher education should teach students how to work critically with AI systems rather than simply prohibiting the tools. It also aligns with Stojanov et al.'s (2024) finding that heavy ChatGPT reliance is associated with weaker independent study habits.

Ethical awareness emerged as another important finding. Students were not uncritical of AI outputs; several reported verifying AI-generated information through textbooks, Google searches, or repeated prompting. They also drew ethical boundaries between acceptable and unacceptable AI use. This suggests that HUFLIT students possess an emerging sense of AI-related academic ethics, though this awareness remains informal. Without institutional guidance, students may develop inconsistent standards of appropriate use. Perkins (2023) similarly argued that institutions need explicit policies addressing AI-assisted work.

Finally, the findings highlight the enduring role of faculty. Despite frequent AI use, students placed greater trust in lecturers' practical experience, pedagogical judgment, and ability to contextualize knowledge. This challenges the assumption that AI will replace human instruction and suggests that faculty roles should be redefined rather than diminished. Lecturers are needed to help students evaluate AI outputs, design process-oriented assessments, and distinguish between perceived learning gains and AI-supported task completion.

In summary, GenAI adoption among HUFLIT students is best understood as pressure-driven, context-sensitive, and pedagogically ambivalent. The extended UTAUT model accurately explains AI use in this setting: while technology features (Effort Expectancy and Performance Expectancy) strongly drive usage, institutional and psychological pressures surrounding graduation certification are equally important. While AI use is associated with students' confidence and perceived readiness, the subjective nature of Perceived Effectiveness demands caution due to the risks of dependency, ethical ambiguity, and a false sense of competence.

7. Conclusion

This study examined how GenAI adoption works among undergraduates preparing for mandatory graduation certificates at HUFLIT. By extending the UTAUT framework with Academic Pressure, the study developed a more context-sensitive model for explaining AI Usage Behavior in a high-stakes educational environment. The integration of quantitative and qualitative findings demonstrates that students' AI adoption is driven by perceived ease of use and usefulness, and is also strongly shaped by systemic academic pressure-particularly the urgency to complete requirements such as TOEIC, VSTEP, MOS, and second foreign language certifications. Effort Expectancy was the strongest direct predictor of AI Usage Behavior, followed by Academic Pressure and Performance Expectancy.

The findings reveal a pedagogical paradox. On the one hand, students perceive AI as effective because it saves time, generates practice materials, clarifies difficult content, and increases confidence. On the other hand, the qualitative interviews indicate that this perceived

effectiveness may sometimes reflect a false sense of competence: students feel capable while using AI but may struggle to internalize the knowledge required for independent performance. AI-assisted learning can therefore function both as a productive scaffold and as a mechanism of cognitive dependency, depending on how students use it. The study contributes theoretically by validating Academic Pressure as a meaningful extension of UTAUT and practically by highlighting the need for universities to guide students toward responsible, reflective AI use.

Practical Implications

A purely prohibitive stance toward GenAI is unlikely to be effective, especially when students already rely on these tools daily. Instead, universities should develop a contextualized "AI-Integrated Academic Integrity Framework." The following strategies are recommended:

1. **Guide AI-assisted Revision:** Provide clear directives on using AI for exam preparation (e.g., generating mock questions, explaining grammar) without substituting independent critical thinking.

2. **Implement Process Documentation:** Require students to maintain and submit prompt logs or process journals when AI is utilized for major assignments, shifting the focus from the final product to the learning journey.

3. **Formalize AI Literacy:** Integrate specific AI literacy modules into foundational academic skills courses to build critical evaluation skills early on.

4. **Redesign Assessment Methods**: Gradually transition toward process-oriented, reflective, and dialogic forms of evaluation, such as oral defenses, in-class tasks, and reflective writing, which require students to explain their thinking process and show understanding without relying only on AI-generated outputs.

5. **Conduct Faculty Training:** Provide targeted professional development for educators focused on AI verification, hallucination detection, ethical citation practices, and strategies for designing AI-resilient curricula.

Limitations and Directions for Further Studies

This study has several limitations. First, it was conducted within a single institution (HUFLIT), and findings may not generalize to other universities, disciplines, or educational systems. Future studies should expand the sample to include students from different Vietnamese institutions-public and private-and from non-language majors to compare how academic pressure and AI usage vary across contexts.

Second, the cross-sectional design captures perceptions and behaviors at one point in time. It cannot determine the long-term effects of AI usage on learning retention, cognitive development, or actual exam performance. Future research should adopt longitudinal designs to examine how frequent AI use relates to students' independent competence over time.

Third, Perceived Effectiveness relied on self-reported perceptions, which may be influenced by social desirability, overconfidence, or the false sense of competence identified in the qualitative findings. Future studies should incorporate objective performance-based measures-pre-tests and post-tests, official certification scores, writing assessments, speaking tasks, or delayed retention tests-to compare perceived effectiveness with actual learning outcomes.

Fourth, the qualitative phase included only six interviewees. Future research should conduct larger-scale qualitative investigations or focus group discussions to explore more

diverse student experiences. Comparing different AI user profiles-highly dependent users, strategic users, skeptical users, and low-frequency users-would help clarify how different engagement patterns affect learning quality.

Finally, future research should investigate the role of lecturers and institutional policy more directly. Studies could examine lecturers' perceptions of AI-assisted certification preparation, their readiness to teach AI literacy, and their strategies for designing AI-resilient assessments. Experimental studies could test the effectiveness of structured AI literacy interventions, comparing students who receive formal training with those who use AI without guidance.

Acknowledgments

We express our sincere gratitude to the Board of Presidents and the Faculty of Foreign Languages at HUFLIT for their support. We are especially thankful to our supervisor, Dr. Tran Quang Bao Phuc, for his valuable guidance, feedback, and encouragement throughout this study. We also deeply appreciate the HUFLIT English Language students who participated in the survey and interviews. Their time and honest responses made this research possible.

References

Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology, 30*(1), 47–88. https://doi.org/10.1111/j.1745-9125.1992.tb01093.x

Ali, R., Tang, O. Y., Connolly, I. D., Fridley, J. S., Shin, J. H., Sullivan, P. L. Z., Doberstein, C. E., Telfeian, A. E., Gokaslan, Z. L., & Asaad, W. F. (2023). Performance of ChatGPT, GPT-4, and Google Bard on a neurosurgery oral boards preparation question bank. *medRxiv*. https://doi.org/10.1101/2023.04.06.23288265

Almogren, A. S., Al-Rahmi, W. M., & Dahri, N. A. (2024). Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon, 10*(11), Article e31887. https://doi.org/10.1016/j.heliyon.2024.e31887

Anh-Hoang, D., Tran, V., & Nguyen, L.-M. (2025). Survey and analysis of hallucinations in large language models: Attribution to prompting strategies or model behavior. *Frontiers in Artificial Intelligence, 8*, Article 1622292. https://doi.org/10.3389/frai.2025.1622292

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology, 54*(5), 1160–1173. https://doi.org/10.1111/bjet.13337

Beyersdorf, K., Klee, S., & Mántica, M. G. (2025). Let's get ethical: An exploration of undergraduate decision making regarding the use of generative artificial intelligence. *Journal of the Student Personnel Association at Indiana University, 52*, 68–89.

Chan, C. K. Y., & Tsi, L. H. Y. (2023). *The AI revolution in education: Will AI replace or assist teachers in higher education?* arXiv. https://doi.org/10.48550/arXiv.2305.01185

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Dang, V. E., Nguyen, D. L. P., & Nguyen, T. H. (2024). Thực trạng ứng dụng ChatGPT trong việc học tập, nghiên cứu của sinh viên Đại học Quốc gia Thành phố Hồ Chí Minh. *Tạp chí Giáo dục, 24*(1), 36–41.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jakkaew, P., & Hemrungrote, S. (2017). The use of UTAUT2 model for understanding student perceptions using Google Classroom: A case study of Introduction to Information Technology course. In *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)* (pp. 205–209). IEEE. https://doi.org/10.1109/ICDAMT.2017.7904962

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Klitgaard, R. (2025). Generative artificial intelligence and universities. *Bhutan Journal of Research and Development, 14*(2), 85–97. https://doi.org/10.17102/bjrd.rub.14.2.005

Ministry of Education and Training. (2017). *Circular No. 12/2017/TT-BGDĐT*. Hanoi, Vietnam.

Nguyen, A. H. (2025). Ảnh hưởng của áp lực học tập đến hành vi sử dụng AI và nguy cơ vi phạm liêm chính học thuật. *Tạp chí Khoa học Trường Đại học Mở Hà Nội, 116*. https://doi.org/10.59266/houjs.2025.670

Nguyen, N. T. (2025). Ảnh hưởng của các đặc điểm trí tuệ nhân tạo (AI) đối với nhận thức hiệu quả học tập của sinh viên đại học. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*. https://doi.org/10.59266/houjs.2025.804

Parveen, K., Tran Quang, B. P., Alghamdi, A. A., Hajjej, F., Obidallah, W. J., Alduraywish, Y. A., & Shafiq, M. (2024). Unraveling the dynamics of ChatGPT adoption and utilization through Structural Equation Modeling. *Scientific Reports, 14*(1), Article 23469. https://doi.org/10.1038/s41598-024-74406-4

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice, 20*(2). https://doi.org/10.53761/1.20.02.07

Pham, T. T. (2024). Đánh giá kết quả phản hồi của Chatbot Bard trong thực hiện bài thi môn Địa lí kì thi tốt nghiệp trung học phổ thông quốc gia Việt Nam từ năm 2019 đến năm 2023. *Tạp chí Giáo dục, 24*(4), 35–40.

Phan, H. H., & Doan, H. C. (2019). Dự báo tác động của Công nghiệp 4.0 đến việc làm của sinh viên sau khi tốt nghiệp. *Tạp chí Khoa học và Công nghệ – IUH, 41*(5). https://doi.org/10.46242/jst-iuh.v41i05.264

Reich, J. (2020). *Failure to disrupt: Why technology alone can't transform education*. Harvard University Press. https://doi.org/10.4159/9780674249684

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing, 53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Stojanov, A., Liu, Q., & Ling, K. J. (2024). University students' self-reported reliance on ChatGPT for learning: A latent profile analysis. *Computers and Education: Artificial Intelligence, 6*, Article 100243. https://doi.org/10.1016/j.caeai.2024.100243

Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education, 49*(2), 223–245. https://doi.org/10.1007/s10755-023-09686-1

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Vo, L. H. P. (2025). Vai trò của trí tuệ nhân tạo (AI) trong giảng dạy bằng tiếng Anh: Nhận thức của sinh viên và ý nghĩa sư phạm cho giảng viên. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*. https://doi.org/10.59266/houjs.2025.778

Wu, Y.-C., Wu, Y.-C., Chang, Y.-C., Yu, C.-Y., Wu, C.-L., & Sung, W.-W. (2025). Advancing medical AI: GPT-4 and GPT-4o surpass GPT-3.5 in Taiwanese medical licensing exams. *PLOS ONE, 20*(6), Article e0324841. https://doi.org/10.1371/journal.pone.0324841

Bionote

Quốc Bảo Mai is an undergraduate student in the Department of English Language, Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT), Vietnam. His research interests encompass educational technology, second language acquisition, and the integration of Generative AI in higher education.

Thị Cẩm Vân Nguyễn is an undergraduate student majoring in English Language at HUFLIT. Her academic pursuits focus on applied linguistics, digital pedagogy, and the psychological impacts of technology-enhanced learning environments on ESL students.

Đình Bảo Hân Nguyễn is an undergraduate student at the Faculty of Foreign Languages, HUFLIT. Her research endeavors primarily revolve around modern pedagogical strategies, technology-enhanced language learning, and the ethical implications of artificial intelligence in academic assessments.

Nguyễn Nhựt Tuyết Hồ is an undergraduate student pursuing a degree in English Language at HUFLIT. She is deeply interested in the intersection of artificial intelligence and language education, specifically exploring behavioral adoption and autonomous learning methodologies among university cohorts.

#176 EQUITY, INCLUSION, AND ACCESS IN HIGHER EDUCATION: CHALLENGES FACED BY STUDENTS FROM DIVERSE SOCIOECONOMIC BACKGROUNDS

Trang Nguyen [1], Vy Nguyen[2]

Abstract

Equity, inclusion and access are fundamental pillars in education, particularly within tertiary education settings, hosting students from diverse socioeconomic backgrounds. his study examines the systemic and classroom-level challenges related to these principles at two public universities in Vietnam: the University of Social Sciences and Humanities (USSH) and Ho Chi Minh City University of Technology (HCMUT). Employing a convergent mixed-methods design, the study utilized an online survey with 100 undergraduate students (50 from USSH and 50 from HCMUT) and semi-structured focus group interviews with 10 participants to investigate students' access to academic resources, perceptions of equity in learning opportunities, and experiences of classroom inclusion. Key findings reveal significant disparities: students from lower-income backgrounds reported financial pressure from hidden financial costs (textbooks, devices, internet, software), digital access barriers, reduced study time due to part-time work, feelings of classroom alienation, and perceptions of insufficient institutional support. These challenges shaped students' everyday learning experiences post-enrollment. The paper proposes evidence-based pedagogical strategies and institutional policy recommendations to promote more equitable and inclusive classrooms. The study also acknowledges its limitations regarding sample representativeness and offers suggestions for future research.

***Keywords*:** equity, inclusion, access, socioeconomic background, higher education

1. Introduction

Higher education is a critical pathway for academic and professional development. However, students do not always experience university life under the same conditions. While pedagogical research often focuses on teaching methods and learning outcomes, less attention has been paid to how socioeconomic backgrounds influence students' ability to participate fully in educational opportunities after enrollment. For many students, the costs associated with learning extend beyond tuition fees and include textbooks, technological devices, internet access, and extracurricular activities. Particularly in Vietnam, where university students come from diverse socioeconomic backgrounds, access to equitable learning experiences sometimes depends on circumstances beyond their control. This difference, however minor, may lead to classroom exclusion rather than opportunity, where a student's academic success may be determined as much by their financial situation as by their individual efforts to acquire knowledge. Hence, the concepts of equity, inclusion, and access are essential in modern

[1]*Nguyen Thi Thao Trang, Foreign Language Center – HCMC University of Technology, VNU-HCM, Ho Chi Minh City, Vietnam Vietnam. Email:trangnguyen@hcmut.edu.vn*

[2]*Nguyen Thi Tuong Vy, English Language Teaching, English Linguistics and Literature, University of Social Sciences and Humanities, VNU-HCM, Ho Chi Minh City, Vietnam*

tertiary education. Although these concepts are widely recognized in educational theory, their practical application is often complicated by the socioeconomic realities of different student groups. For this reason, understanding how socioeconomic backgrounds shape students' access to resources, their sense of inclusion, and their experience of equity within the learning environment has become a crucial concern in Vietnam's higher education system.

This study contributes to the body of existing research by redirecting the focus from the general academic outcomes or the comprehensive access to higher education to the daily lived experiences of socioeconomic inequality that students encounter after enrolling. It investigates not just whether students join university, but how they experience equity, access, and inclusion during their university education. The research was conducted to address the following question:

In which way do Ho Chi Minh City University of Technology (HCMUT) and University of Social Sciences and Humanities (USSH) students perceive the challenges regarding equity, access and inclusion based on their diverse socioeconomic backgrounds?

2. Literature Review

2.1. Equity, Inclusion, and Access in education

Modern education has to be fundamentally restructured in light of the rapidly changing student demographics brought about by social, political, and cultural shifts. In 21st-century education, to avoid classroom alienation and systemic inequality, the core values of equity, access, and inclusion are no longer optional or theoretical ideas; they are essential for constructing equitable, sustainable, and forward-thinking learning environments (Panigrahi, 2025). To provide a theoretical foundation for the survey instruments used in this study, this section reviews literature surrounding three main constructs: equity, access, and classroom inclusion in education.

According to the Strength through Diversity Project (OECD, 2023), equity in education systems means that all students can achieve their educational potential, regardless of their gender, ethnic origin, Indigenous background, immigrant status, sexual orientation and gender identity, special education needs, or giftedness (Cerna et al., 2021; OECD, 2017). To understand this concept, it is essential to distinguish between equality and equity. Panigrahi (2025) and Iheduru-Anderson et al. (2025) clarified that equity is different from equality in that, whereas equality just means the same treatment for all individuals, equity acknowledges the existence of systemic imbalances and demands interventions to address them. Yasinzai and Ahad (2025) further noted that achieving fairness in education necessitated acknowledging and resolving historical and systemic disadvantages.

Iheduru-Anderson et al. (2025) defined access as the elimination of structural, economical, physical, technical, systemic, curricular, language, and cultural barriers to full participation in educational and professional settings via appropriate accommodations. Although worldwide higher education enrollment has increased, access remains significantly uneven across different regions and social and economic classes (Yasinzai & Ahad, 2025). With a growing dependency on online education platforms, Panigrahi (2025) pointed out that it is critical to ensure educational materials and resources are accessible to students with learning disabilities and those living in distant areas. To him, accessibility also applies to language, pedagogy, and evaluations, ensuring that students may participate meaningfully regardless of where they begin or their personal background.

UNESCO (2009) characterized inclusive education as "an on-going process aimed at offering quality education for all while respecting diversity and the different needs and abilities, characteristics and learning expectations of the students and communities, eliminating all forms of discrimination". The goal for it is to address all students' needs, which extends beyond school attendance and achievement, while simultaneously improving their well-being and participation (Cerna et al., 2021), so that diverse individuals can feel valued, welcomed, respected, and able to contribute fully within an environment (Panigrahi, 2025).

Table 1

Defining Equity, Access, and Inclusion

Terms	Definition
Equity	Fairness in treatment, access, and outcomes by addressing structural and systemic barriers.
Access	Removing structural, cultural, institutional, and systemic barriers that limit full participation or advancement.
Inclusion	Creating environments where all individuals feel welcomed, valued, respected, and able to participate and contribute fully.

**Note.* Adapted from Iheduru-Anderson et al.(2025)

2.2. Socioeconomic backgrounds and inequality in higher education

These theoretical distinctions regarding the three concepts are crucial, as socioeconomic backgrounds play an essential role in determining learning opportunities within higher education. As stated by Tomaszewski et al. (2022), people from socioeconomically deprived backgrounds frequently have lower academic performance than those from more socio-economically advantaged backgrounds; this disparity can be linked to their fundamental differences in learning experiences.

The claim was also further supported by previous studies from Sabic-El-Rayess (2012, 2014 & 2020); Sabic-El-Rayess & Mansur (2016), Sabic-El-Rayess & Seeman (2017), Sabic-El-Rayess & Otgonlkhagva (2012), Sabic-El-Rayess, Mansur, Batkhuyag & Otgonlkhagva (2019), Moratti and Sabic-El-Rayess (2009), who demonstrated that elite mechanisms and interests often prevail over the needs of the disadvantaged across social, economic, and educational hierarchies globally. Unfortunately, universities and colleges only help reinforce this existing socioeconomic strata further (Sabic-El-Rayess & Mansur, 2016), unless there is a strong political determination to disrupt the status quo and increase opportunities for the poor (Fadhil & Sabic-El-Rayess, 2021). As a result, students from disadvantaged backgrounds, such as those with disabilities, those living in poverty, ethnic and linguistic minorities, and rural communities, continue to encounter systemic challenges that impede their ability to study, contribute, and achieve (Panigrahi, 2025).

According to Tomaszewski et al. (2022), there are seven different barriers to accessing university, such as low self-efficacy, received youth allowance, low-SES school, school life, no plan for university, parents not wanting university and low PISA (Programme for International Student Assessment) score. The findings in their research showed that individuals

from lower socioeconomic origins are far more likely than their more advantaged counterparts to encounter these numerous challenges to attending college.

Applied to our study, these findings suggest that simply providing the same course requirements to every student is insufficient. When learners lack access to standard requirements for higher education learning such as the financial means to purchase coursebooks or afford tuition fees, the "equal" treatment often masks an underlying inequality.

2.3. The Vietnamese higher education context

Many studies have been undertaken on inequities in educational access in Vietnam. Nguyen (2004) observed that while countrywide school enrollment expanded across all educational levels, particularly in higher education, the financial burden of private schooling heavily impacted low-income households. This economic strain further widened the disparity between affluent and disadvantaged students regarding both educational opportunity and quality. Another study by Thach (2020) indicated that children from low-income families' access to higher education lower than their peers. According to Vu (2012), severe inequalities in school access still exist; hence, significant gaps continue to separate rural and urban students, ethnic minorities and the majority population, and wealthy households from low-income ones.

To operationalize these concepts for the current study, the literature is synthesized around three main constructs that directly inform our survey and interview instruments:

1. Access to resources: focused on structural and material barriers (Iheduru-Anderson et al., 2025), this construct measures students' ability to afford hidden costs such as laptops, internet, textbooks and proprietary software.

2. Perception of equity: rooted in the acknowledgment of systemic imbalances (Panigrahi, 2025), this construct evaluates the perceived fairness of university support systems and rigid academic policies.

3. Perception of inclusion: based on UNESCO's (2009) framework, this construct assesses students' psychological sense of belonging and their comfort in participating without fear of socioeconomic alienation.

2.4. Research gap

While existing literature highlights the challenges low-socioeconomic status imposes on higher education students across the globe, there remains a notable lack of contemporary research specifically situated within the Vietnamese higher education context. This study seeks to address this and bridges three specific gaps. First, a contextual gap exists as most prior studies focus on enrollment or broad academic outcomes rather than the everyday classroom experiences in Vietnamese public universities. Second, a conceptual gap remains regarding the intersection of the three pillars of modern education: equity, access, and inclusion; few studies examine how they influence one another simultaneously. Third, there is an empirical gap in qualitative exploration into students' lived experiences of hidden academic costs and classroom alienation post-enrollment. This study seeks to close these gaps by exploring how these three pillars influence one another, and how they appear differently for students' daily experiences from diverse socioeconomic backgrounds at two Vietnamese public universities (USSH and HCMUT). By connecting specific barriers that hinder equitable, accessible, and inclusive learning experiences, this research proposes pedagogical strategies and institutional policy recommendations for a just and sustainable educational framework that ensures all learners can achieve their full academic potential, regardless of their socioeconomic backgrounds.

3. Methodology

3.1. Research site

The research was conducted at two member institutions of Vietnam National University Ho Chi Minh City (VNU-HCM): the University of Social Sciences and Humanities (USSH) and Ho Chi Minh City University of Technology (HCMUT). USSH is located in the VNU-HCM Urban Area, Linh Trung Ward, Thu Duc City, Ho Chi Minh City, while HCMUT's second campus is located in Dong Hoa Ward, Ho Chi Minh City. As two major public universities in southern Vietnam, these institutions attract students from diverse socioeconomic and educational backgrounds across the country. USSH and HCMUT were treated as two sites for exploring shared student experiences rather than as a comparative study; no institutional differences were systematically examined.

3.2. Participants

The participants were undergraduate students from the Faculty of English Linguistics and Literature (EF) and Engineering Faculty. A total of 100 students agreed to take part in the study, with 50 students from USSH and 50 students from HCMUT. The convenience sampling approach was implemented to select the sample. This approach was chosen due to the researchers' existing connections within the faculties, which is efficient and provides timely data collection. By targeting students based on their availability and willingness to participate, the researchers were able to gather a diverse range of student perspectives regarding academic challenges.

However, this non-probability sampling limits the full representativeness of the findings, meaning the findings cannot be generalized to all students at these universities or Vietnamese higher education broadly, as participation was voluntary and dependent on accessibility. Participant variation included multiple academic years (freshmen to seniors), gender, major, and self-reported family financial status (low income, lower-middle income, middle income, upper-middle income, high income). Detailed data on rural/urban background, scholarship status, and living arrangements (with family or renting) were also not systematically collected; future studies should include these as socioeconomic barriers may differ across conditions.

3.3. Study design

This study employed a convergent mixed-methods design (Creswell & Plano Clark, 2018). Quantitative and qualitative data were collected concurrently during the same period to provide a comprehensive understanding: the survey (quantitative data) captured broad patterns of perceptions, while follow-up focus group interviews (qualitative data) provided deeper explanations of students' lived experiences. The two strands were integrated during interpretation to complement each other.

3.4. Data collection procedures

For quantitative data, an online questionnaire was distributed via email to 100 participants. The items from the survey were directly developed based on the conceptual definitions outlined in the literature review (Table 1). It consisted of 18 Likert-scale items (1 = Strongly Disagree to 5 = Strongly Agree) and 3 open-ended questions, covering five key sections aligned with the three core constructs:

1. Access to resources: Items derived from structural/material barriers in literature included affordability of textbooks, devices, internet, software, and extracurricular costs.

2. Perception of equity: Items derived from fairness and unequal resources focused on fairness of assessments, influence of external resources, and university support.

3. Perception of classroom inclusion: Items derived from belonging and participation evaluated the sense of belonging, discomfort sharing experiences, and social alienation.

For qualitative data, semi-structured focus group interviews (10 questions) with 10 participants served as follow-up to explore hidden costs, classroom exclusion, policy flexibility, and institutional support in depth. Interviews were conducted in familiar classroom settings, audio-recorded with consent, and supplemented with observation notes.

3.5. Reliability and validity

A pilot test of the questionnaire was conducted with 15 students from similar backgrounds to assess clarity, face validity, and reliability. Internal consistency reliability was evaluated using Cronbach's alpha for each subscale: access to resources: $\alpha = 0.82$; perception of equity: $\alpha = 0.79$; perception of classroom inclusion: $\alpha = 0.85$. All coefficients exceeded the acceptable threshold of 0.70 recommended for exploratory research in social sciences (Hair et al., 2019), indicating internal consistency. Content validity was further supported through instrument alignment with established definitions and consultation from higher education faculty members.

Regarding the qualitative data, transcript coding was independently conducted and subsequently cross-checked by two researchers to resolve discrepancies and establish consensus, ensuring trustworthiness.

3.6. Data processing procedures

An online questionnaire was distributed via email. Raw data from the questionnaires were processed using the Statistical Package for the Social Sciences (SPSS). The results were extracted and then analyzed using frequency tables and bar charts. Qualitative data from the focus group interviews were transcribed verbatim and analyzed using thematic analysis following Braun and Clarke's (2006) six-phase approach. An inductive coding process was employed, while also being informed deductively by the three core constructs (access to resources, perceptions of equity, and classroom inclusion) from the literature and survey. Initial codes were generated, collated into potential themes, reviewed, defined, and named. To enhance trustworthiness, two researchers independently coded a subset of transcripts, and discrepancies were resolved through discussion until consensus was reached. Member checking was conducted with two participants to verify the accuracy of interpreted themes.

3.7. Ethical issues

Since the research involves audio recordings, semi-structured focus group interviews, and questionnaires, we ensure that informed consent is obtained from all participants prior to their involvement. For the questionnaires, we sent an invitation via email that included all the necessary details about the study, including its benefits and any possible risks. Before the participants filled out the form, they had to sign a consent statement, which confirmed they were aware of the points and allowed the research to use their data. Regarding the interview, the participants are only asked to discuss topics they are comfortable with and they will have the right to refuse to participate in the research or withdraw from it at any time. To ensure confidentiality in the results, all names were replaced with numbers (e.g., Participant 01). Hard copies of data were stored in secured drawers, and digital files were password-protected.

4. Results

After reviewing the survey responses and conducting focus group interviews, the researchers processed 100 survey questionnaire responses and 10 in-depth interview transcripts.

The quantitative analysis reveals a significant disparity in how students perceive equity, resource access and inclusion at the university. Table 2 shows that most respondents came from middle-income (40%) and lower-middle-income (25%) backgrounds. In addition, 62% of participants reported not having a part-time job, while 38% were engaged in part-time employment.

Table 2

Demographic Information of the Respondents

Variable	Category	Number of students	Percentage
Family financial status	Low income	15	15%
	Lower-middle income	25	25%
	Middle income	40	40%
	Upper-middle income	15	15%
	High income	5	5%
Part-time job	Yes	38	38%
	No	62	62%

**Note.* Data are based on student responses to survey questions regarding family financial status and part-time employment status. $N = 100$.

The qualitative investigation, conducted through focus group interviews with 10 survey respondents, provides deeper insights into the lived experiences and challenges students encounter.

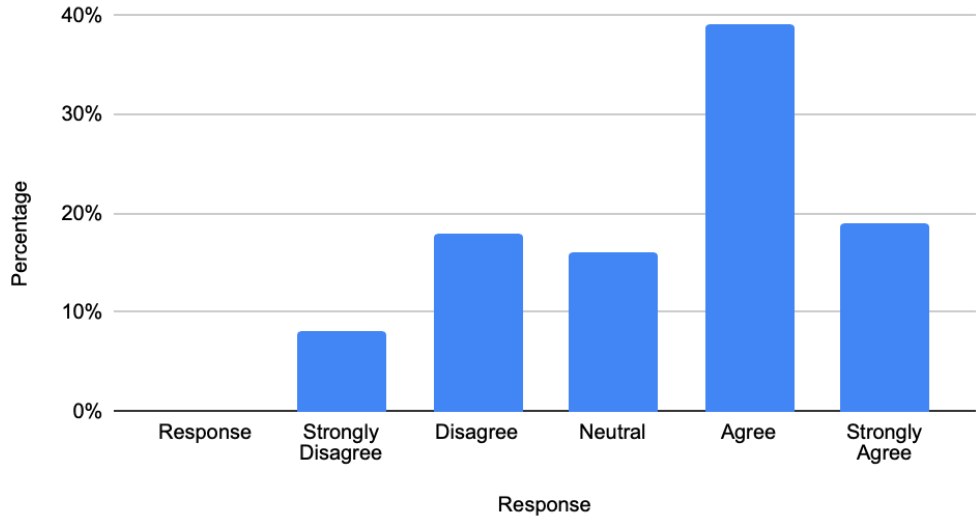
To directly address the research questions, the findings regarding students' learning experiences are organized conceptually around the three core pillars: access to resources, perceptions of equity, and classroom inclusion, integrating both quantitative trends and qualitative depths.

4.1. Access to resources

Figure 1 illustrates students' perceptions regarding their ability to afford necessary learning resources. While 39% of respondents agreed that they were able to afford the required resources, 26% of students expressed concerns regarding the burden of academic requirements (Mean = 3.21). This suggests that the costs of textbooks, internet access, software, technological devices and other learning expenses remain barriers for a notable minority of students.

Figure 1

Student Responses Regarding Financial Ability to Afford Necessary Learning Resources

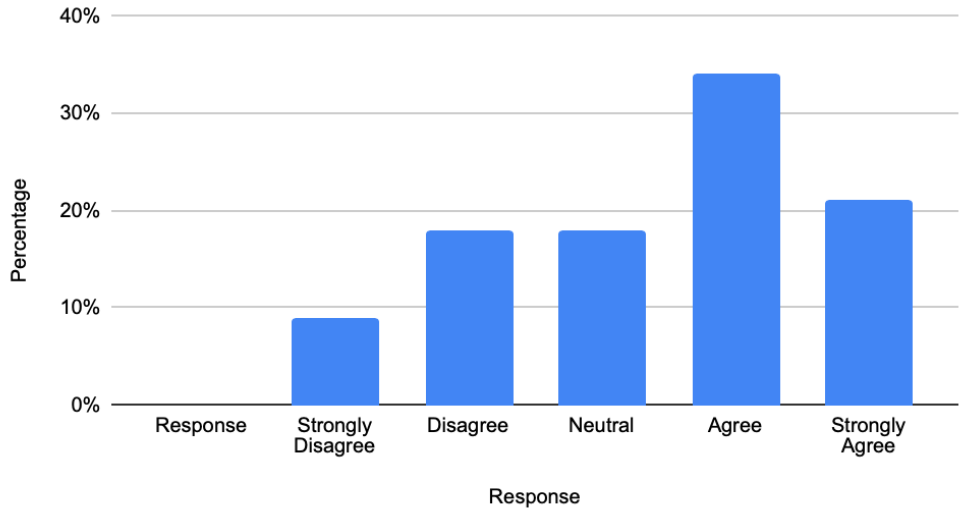


Note. Data are based on student responses to the survey item: "My financial situation allows me to afford all necessary learning resources (e.g., textbooks, specialized software, learning applications, a stable internet connection) and other course-related expenses".

Figure 2 presents students' responses regarding the impact of work responsibilities on their studies. Approximately 34% of respondents agreed and 21% strongly agreed that work responsibilities reduced the time and energy they could devote to academic activities. This suggests that balancing work and study can be a challenge for many students.

Figure 2

Students' Perceptions of the Impact of Work Responsibilities on Study Time and Energy



Note. Data are based on student responses to the survey item: "My work responsibilities (eg house chores, part-time jobs) reduce the time and energy I can devote to my studies".

These statistical patterns were deeply echoed in the focus group interviews as interviewed students expressed severe frustration regarding the "hidden costs of digital and academic learning". 6 out of 10 students noted that these requirements are often treated as standard, overlooking the financial strain on low-income students.

"For our speech and presentation classes, we are constantly required to create high-quality slides and digital portfolios. My laptop is a second-hand model my family bought years ago so it runs incredibly slow and the battery dies in 20 minutes. When professors require us to buy original international textbooks instead of using photocopies due to copyright rules, it takes a massive chunk out of my monthly allowance. I ended up having to borrow a friend's book to take photos of the pages." - Participant 03.

4.2. Perception of equity

While some students acknowledge the support of the university's support services such as library and computer labs, respondents with lower-income status viewed current university support as insufficient in addressing their needs. Table 3 shows that when assessing proactive institutional support, disparities emerged across income groups. 60% of students with low income, 40% of lower-middle income and 22% of middle income perceived the university support services as inadequate. Although these percentages highlight a stark contrast based on self-reported financial status, it is important to note that inferential statistical tests (e.g., chi-square) were not conducted; thus, these figures represent observed trends within the convenience sample rather than statistically generalized differences.

Table 3

Students' Perceptions of Proactive University Support by Financial Status

Financial status	Agree/Strongly agree	Neutral	Disagree/Strongly disagree
Low income (15)	3 (20%)	3 (20%)	9 (60%)
Lower-middle income (25)	8 (32%)	7 (28%)	10 (40%)
Middle income (40)	19 (48%)	12 (30%)	9 (22%)
Upper-middle income (15)	9 (60%)	4 (27%)	2 (13%)
High income (5)	3 (60%)	2 (40%)	0 (0%)

**Note.* Data are based on student responses to the survey item: "I feel the university takes proactive steps to support students from low-income backgrounds (e.g., providing facilities and scholarships) so that they are not left behind".

Qualitatively, in the interviews, 5 out of 10 students also emphasized the severe stress of balancing part-time employment with academic deadlines. The strict attendance and submission policies at their universities were cited as lacking the flexibility needed for working students.

"I work as a barista at a coffee shop and do evening tutoring to cover my room rent in the city. My shifts sometimes conflict with morning classes if a coworker gets sick. The university's attendance policy is incredibly rigid; if you miss more than 20% of classes, you are barred from taking the final exam. If I am late because a shift ran over, my process score suffers." - Participant 07.

Moreover, 6 out of 10 participants expressed concerns that the university does not truly understand the extent of the equity gap. 2 out of 10 participants felt that policies are designed for the students from more affluent backgrounds, ignoring those facing financial hardship. 3 out of 10 participants also noted that the university's support systems were often reactive rather than preventive. Assistance was usually available only after students had already encountered difficulties. However, 3 out of 10 participants noted that some instructors show genuine empathy regarding socioeconomic challenges. This includes providing recorded lectures or accepting late submissions due to unavoidable work shifts.

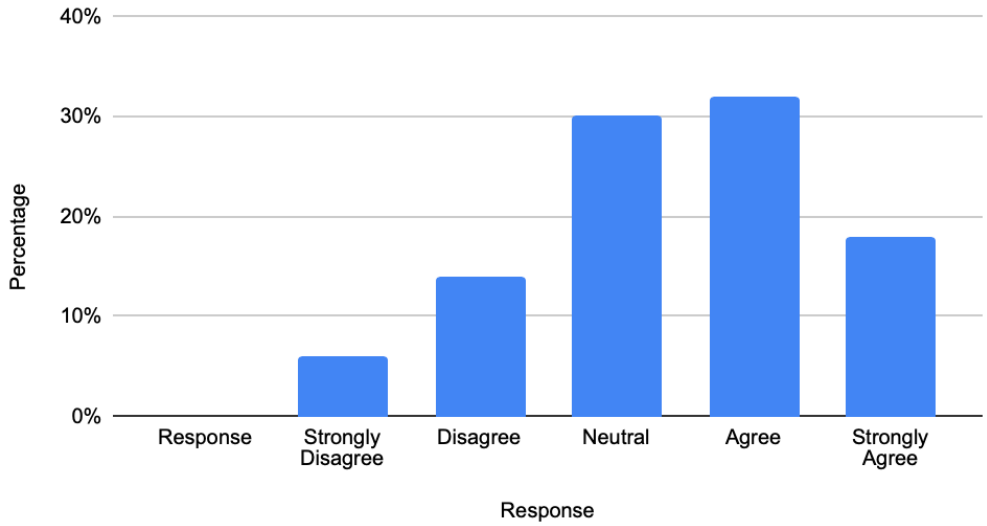
"The school's administrative system is slow and the computer labs are outdated, which doesn't help much. But some of my instructors have been amazing. When I explained that I didn't have Wi-Fi at home last month, my teacher allowed me to submit my essay a day late." - Participant 09.

4.3. Perception of inclusion

Figure 3 illustrates students' perceptions of inclusivity in the classroom. A notable proportion of respondents (32%) reported feeling uncomfortable sharing personal experiences that differed from those of wealthier peers. This finding suggests that socioeconomic differences may influence students' sense of inclusion and participation in classroom discussions.

Figure 3

Students' Feelings of Discomfort Sharing Personal Experiences due to Socioeconomic Differences



Note. Data are based on student responses to the survey item: "I feel uncomfortable sharing my personal experiences in class as they differ from the "typical" or "wealthy" student experience".

This sense of subtle exclusion was detailed in the qualitative findings. A recurring theme among interview participants was the feeling of being "out of place" within certain classroom discussions and social interactions. Most students from lower-income backgrounds reported feeling left out when topics related to travel, private tutoring, expensive hobbies, or dining experiences were discussed in class. Although these conversations were not intended to exclude anyone, 3 out of 10 participants noted that they often struggled to relate to such experiences. As a result, they became less willing to participate in discussions or share their own personal experiences, fearing that they would be judged or misunderstood by their peers.

5. Discussion

The findings suggest that socioeconomic differences not only appear to prevent students from entering universities but also continue to influence how students experience university life after enrollment. This extends prior research (e.g., Tomaszewski et al., 2022; Panigrahi, 2025;) by revealing "post-access" barriers inside higher education as challenges that persist even after enrollment and affect equity, access, and inclusion simultaneously.

Access to education therefore extends beyond admission and tuition fees. Reliable internet access, appropriate learning devices, academic materials, and the ability to participate in extracurricular activities all contribute to students' opportunities to engage fully in their studies. For many respondents with stronger financial backgrounds, these requirements are often treated as a standard expectation. For other notable numbers with lower-socioeconomic backgrounds, they represent additional obstacles that must be overcome alongside academic responsibilities.

The results also raise questions about educational equity. Students in the same classrooms are required to reach the same academic standards, but they do not have the same resources or opportunities. Lower-income students were more likely to regard university support as insufficient and to report problems due to financial pressures. At the same time, students who worked part-time frequently described challenges balancing employment and academic responsibilities. Equal expectations may therefore not always lead to equal opportunities. When students must devote significant time and energy to meeting financial needs, their ability to participate in learning activities can be reduced despite having similar academic potential.

Inclusion appears as a more subtle issue. Within this sample, most participants reported a general sense of belonging within their faculties; however, feelings of exclusion often appeared during everyday classroom interactions rather than through formal institutional practices. Discussions involving travel experiences, private tutoring, or expensive leisure activities occasionally reinforced social differences between students. Such occurrences may seem minor, yet repeated exposure can contribute to feelings of isolation among students from unprivileged backgrounds. This emphasizes the importance of creating learning environments that acknowledge and respect students' diverse socioeconomic experiences.

The results highlight the interconnected nature of the three pillars: limited access undermines equity, which in turn affects inclusion.

It is important to clarify that this study provides perception and experience-based evidence from a small convenience sample at two specific faculties. The findings suggest patterns within this specific sample rather than broad systemic facts, thereby should not be generalized broadly without qualification. USSH and HCMUT were treated as complementary sites for exploring shared experiences rather than compared directly; no institutional differences were tested.

6. Conclusion

This study provides insights into the socioeconomic barriers affecting students at USSH and HCMUT in relation to equity, access, and inclusion in higher education. The findings suggest that access, equity, and inclusion should not be viewed only as separate concepts. Access to educational resources alone does not necessarily guarantee equitable learning opportunities, while equitable opportunities do not automatically ensure that all students feel included within the learning environment. Students from lower socioeconomic backgrounds often faced additional challenges that were less visible to their peers, particularly in relation to educational expenses, work responsibilities, and participation in university life. Efforts to reduce socioeconomic barriers should therefore extend beyond financial assistance and consider how institutional policies, academic expectations, and support services affect students with different levels of economic resources.

Institutional-level recommendations: for digital and material support, universities should subsidize essential digital subscriptions and improve access to updated computer

laboratories to address the technological deficit reported by students lacking adequate devices. To ensure policy flexibility, schools should establish more flexible attendance and assessment deadline policies specifically tailored to accommodate students working part-time, alleviating the severe time pressures identified in the study. Lastly, for proactive communication, communication regarding available financial aid and scholarships should be strengthened so that support systems act preventively rather than reactively.

Teacher-level recommendations: regarding inclusive pedagogy, educators should be trained to avoid making assumptions about students' financial resources, offering low-cost alternatives (e.g., OER materials instead of standard expensive imported textbooks). Teachers must be more sensitive towards classroom inclusivity by actively guiding class discussions to ensure discussing topics (e.g., daily experiences, hobbies) do not alienate students from disadvantaged backgrounds, fostering a culturally and economically sensitive environment.

While this study sheds light on important issues surrounding equity, access, and inclusion, it has several limitations. First, the reliance on a relatively small convenience sample restricts the generalizability of the findings. Second, the data is self-reported, and socioeconomic status was self-categorized by students rather than measured through objective economic indicators (e.g., household income tax or official pell-grant equivalents). Third, utilizing focus groups for sensitive financial discussions may have introduced social desirability bias, as some students might have hesitated to openly share deep financial hardships in front of peers.

Future research should involve a broader range of participants across different institutions and employ longitudinal approaches to track how socioeconomic factors influence students throughout their academic journeys. Furthermore, conducting one-on-one individual interviews is recommended for future qualitative investigations to provide a safer space for discussing sensitive financial topics. Finally, comparative research designs such as analyzing differences between public and private universities, or urban versus rural student populations would yield a more comprehensive understanding of systemic inequities in Vietnamese education. Ultimately, addressing these barriers is essential to ensuring that higher education remains equitable, accessible, and inclusive for students regardless of their socioeconomic backgrounds.

Acknowledgments

The author would like to express sincere gratitude to all participants for their valuable contributions to this study. Special appreciation is extended to the survey respondents for their time, cooperation, and honest feedback, which were essential to ensuring the reliability of the research data. The author also thanks colleagues and mentors for their insightful comments and continuous support throughout the research process. The author gratefully acknowledges the organizing committee of the HUFLIT International Conference for providing a valuable academic platform for knowledge exchange and scholarly discussion.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Q*ualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cerna, L., Mezzanotte, C., Rutigliano, A., Brussino, O., Santiago, P., Borgonovi, F., & Guthrie, C. (2021). *Promoting inclusive education for diverse societies: A conceptual framework* (OECD Education Working Papers No. 260). OECD Publishing. https://doi.org/10.1787/94ab68c6-en

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Fadhil, I., & Sabic-El-Rayess, A. (2021). Providing equity of access to Higher Education in Indonesia: A policy evaluation. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 57–75. https://doi.org/10.23917/ijolae.v3i1.10376

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning. https://www.cengage.uk/c/multivariate-data-analysis-8e-hair-babin-anderson-black/9781473756540/?searchIsbn=9781473756540

Iheduru-Anderson, K., Jayaraman, R., Inungu, J., & Mumford, V. E. (2025). Diversity, equity, inclusion, and access: Confronting the myth of meritocracy in nursing academia. *Electronic Journal of General Medicine*, *22*(6). https://doi.org/10.29333/ejgm/17437

Moratti, M., & Sabic-El-Rayess, A. (2009). *Transitional justice and DDR: The case of Bosnia and Herzegovina* (Research Brief). International Center for Transitional Justice. https://www.ictj.org/sites/default/files/ICTJ-DDR-Bosnia-ResearchBrief-2009-English_0.pdf

Nguyen, N. N. (2004). Trends in the education sector. In P. Glewwe, N. Agrawal, & D. Dollar (Eds.), *Economic Growth, Poverty, and Household Welfare in Vietnam* (pp. 415–438). World Bank.

Organisation for Economic Co-operation and Development. (2017). *Educational opportunity for all: Overcoming inequality throughout the life course* (Educational Research and Innovation). https://doi.org/10.1787/9789264287457-en

Organisation for Economic Co-operation and Development. (2023). *Equity and inclusion in education: Finding strength through diversity*. https://doi.org/10.1787/e9072e21-en

Panigrahi, S. (2025). Equity, inclusion and accessibility in Education. *International Journal For Multidisciplinary Research*, *7*(6). https://doi.org/10.36948/ijfmr.2025.v07i06.64299

Sabic-El-Rayess, A. (2012). *Making of a voiceless youth: Corruption in Bosnia and Herzegovina's higher education* (Doctoral dissertation, Columbia University). Academic Commons. https://doi.org/10.7916/D8K361NJ

Sabic-El-Rayess, A. (2014). Acting and reacting: Youth's behavior in corrupt educational settings. *Peabody Journal of Education*, *89*(1), 70–85. https://doi.org/10.1080/0161956X.2014.862473

Sabic-El-Rayess, A., & Mansur, N. N. (2016). Favor reciprocation theory in education: New corruption typology. *International Journal of Educational Development*, *50*, 20–32. https://doi.org/10.1016/j.ijedudev.2016.04.005

Sabic-El-Rayess, A., Mansur, N. N., Batkhuyag, B., & Otgonlkhagva, S. (2019). School Uniform Policy's adverse impact on equity and access to schooling. *Compare: A Journal of Comparative and International Education*, *50*(8), 1122–1139. https://doi.org/10.1080/03057925.2019.1579637

Sabic-El-Rayess, A., & Otgonlkhagva, S. (2012). *School uniform cost reduction study: Standardization, simplification and supply policy*. International Center for Transitional Justice; Columbia University. https://doi.org/10.7916/d8-6e50-8257

Sabic-El-Rayess, A., & Seeman, A. M. (2017). *America's familial tribalism*. Academic Commons. https://doi.org/10.7916/d8-7jt8-mh87

Thach, T. T. (2020). *Provide equal access to education*. Fulbright University Vietnam. https://fulbright.edu.vn/vi/trao-co-hoi-tiep-can-giao-duc-mot-cach-binh-dang/

Tomaszewski, W., Perales, F., Xiang, N., & Kubler, M. (2022). Differences in higher education access, participation and outcomes by socioeconomic background: A life course perspective. *Life Course Research and Social Policies*, 133–155. https://doi.org/10.1007/978-3-031-12224-8_7

UNESCO (2009). *Defining an Inclusive Education Agenda: Reflections around the 48th session of the International Conference on Education*. http://www.ibe.unesco.org/sites/default/files/resources/defining_inclusive_education_agenda_2009.pdf

Vu, L. H. (2012). An overview of access to and inequality in the education system of Viet Nam. *Asia-Pacific Development Journal*, *19*(1), 37–62. https://doi.org/10.18356/57405f5e-en

Yasinzai, F., & Ahad, A. (2025). Equity, inclusion, and access in higher education: Challenges, progress, and Future Directions. *Research Journal of Psychology*, *3*(4), 292–295. https://doi.org/10.59075/rjs.v3i4.270

Bionote

Nguyen Thi Thao Trang (Ms) holds a Master's degree in TESOL through a joint program with Victoria University, Australia. She has over 16 years of experience teaching Non-English-Major learners at Ho Chi Minh City University of Technology and at Language Centers specializing in IELTS and TOEIC. Her research and teaching interests focus on technology-enhanced language learning, soft skills development, and innovation in English education. She has led student teams in award-winning community projects and participated in international professional development programs, including Train the Trainer and online instructional design workshops. She continues to explore innovative approaches to enhance language learning outcomes and foster learner autonomy in higher education contexts.

Nguyen Thi Tuong Vy (Ms.) holds a Bachelor's degree in English Linguistics and Literature from the University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City. She has four years of experience teaching English at language centers and is currently an English Teacher at RES English Center. Her professional interests include English language teaching and language learning. She can be contacted at nguyengon196@gmail.com.

Appendices

Appendix A – Survey questions

Socioeconomic barriers to equity, access and inclusion in education for students with diverse socioeconomic backgrounds

Student Information

1. Academic year: _____

2. What's your email address? _____

3. Do you work part-time?

1. Yes
2. No

If yes, what is your part-time job? _____

4. How would you describe your family's financial situation?

1. Low income
2. Lower-middle income
3. Middle income
4. Upper-middle income
5. High income

Instructions: *Please indicate your level of agreement using a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.*

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

Questions

Part 1: Perception of socioeconomic impact on access

1. My financial situation allows me to afford all necessary learning resources (e.g., textbooks, specialized software, learning applications, a stable internet connection) and other course-related expenses.
2. The cost of participating in extra-curricular activities (e.g., workshops, certificates) is a burden for me.
3. I have reliable access to a high-performance computer/laptop to complete digital assignments.
4. My home environment (e.g., lack of quiet space, household responsibilities) limits my ability to study outside of class.
5. My work responsibilities (e.g., house chores, part-time jobs) reduce the time and energy I can devote to my studies.

Part 2: Perception of socioeconomic impact on equity

1. I feel that access to external resources (e.g., private tutors, expensive software) influences academic outcomes more than students' learning efforts.

2. Course assessments are designed in a way that is fair to students regardless of their financial background.

3. I feel that students with higher financial resources have an unfair advantage in keeping up with the course pace.

4. The university provides sufficient support (e.g., computer labs, library resources) to help students who lack equipment at home.

Part 3: Perception of socioeconomic impact on inclusivity

1. I feel uncomfortable sharing my personal experiences in class as they differ from the "typical" or "wealthy" student experience.

2. My instructors show awareness of the different socioeconomic challenges students face.

3. The language and topics used in my classes are not relatable to my life experiences and background.

4. I feel a sense of belonging in the faculty, regardless of my socioeconomic status.

Part 4: Perception of systemic challenges

1. Financial concerns negatively affect my participation in activities compared to my peers (e.g., class discussions, group work, workshops).

2. Students who experience financial difficulties can easily access information about available support services at the university (e.g., scholarships, financial aid)

3. I feel the university's policies on attendance and deadlines are flexible enough to accommodate students with part-time jobs or financial difficulties.

4. I feel the university takes proactive steps to support students from low-income backgrounds (e.g., providing facilities and scholarships) so that they are not left behind.

Part 5: Open-ended questions

1. Are there any financial challenges (e.g., learning materials, internet access, devices, transportation, extracurricular costs) that affect your learning experience? Please explain.

2. Have you ever felt "different" or "out of place" in a classroom due to your financial or social background? Describe the situation.

3. What additional support or changes should the university provide to help students from different socioeconomic backgrounds?

Appendix B – Interview questions

1. How does your current financial situation influence your ability to access the tools and resources you need to succeed in your classes?

2. Have you ever felt that your socioeconomic background made it harder for you to participate in or keep up with course requirements compared to your classmates?

3. Do you believe your instructors are aware of and sensitive to the different personal or financial challenges that students face outside of the classroom?

4. Can you describe a time when you felt either included or excluded in a classroom setting based on your social or economic background?

5. What is one practical change you think the University could make to ensure that all students, regardless of their financial status, have an equal chance in the program?

#178 AI-DRIVEN SELF-DIRECTED LANGUAGE LEARNING: STRATEGY USE AND GENDER PERSPECTIVES AMONG VIETNAMESE EFL UNDERGRADUATES

Huynh Thanh Trung [1], Chau Thi Kim Ngan [2], Vo Thi Tien [3], & Tran Quoc Thao [4]

Abstract

Artificial intelligence (AI) has recently emerged as a prominent force in language learning, opening opportunities to foster students' learning autonomy and strengthen their self-directed learning skills. AI is reshaping educational settings at a dramatic pace. Yet evidence on how students in Vietnam actually put these tools to work in their own learning remains limited. This study examined how Vietnamese EFL undergraduates at one vocational college employ AI-driven self-directed language learning strategies (AI-SDLLS), and whether gender shapes strategy use. Adopting a convergent mixed-methods design, it paired a questionnaire (N = 52) with semi-structured interviews (n = 16). The findings revealed a consistent descriptive pattern: cognitive strategies such as grammar correction, translation, and idea generation were used most frequently, metacognitive and affective strategies surfaced at moderate levels, and social strategies ranked lowest. This pattern suggests that students gravitate toward AI for immediate, task-level support but underuse it for planning, self-monitoring, or collaborative learning. As for gender, independent-samples t-tests detected no significant differences across the four subscales, though the modest sample limits statistical power. Given its single-institution scope, the study offers context-specific evidence rather than a national picture. Targeted AI-literacy and metacognitive-strategy instruction, coupled with teacher preparation for responsible integration, could support Vietnamese classrooms as they move toward English as a Second Language (ESL) practice.

***Keywords*:** artificial intelligence, EFL to ESL, gender, self-directed language learning strategies

1. Introduction

Vietnam's English-language landscape is shifting. Driven by national policy and labor-market demand, English is increasingly positioned as a language for higher education, employment, and global participation (Nguyen & Tran, 2025). This repositioning is best understood as a policy and pedagogical aspiration, an ongoing reorientation toward English as a second language (ESL), rather than a completed sociolinguistic shift, since for most Vietnamese learners English remains a foreign language encountered largely in the classroom. Realizing this aspiration depends in part on what happens beyond the classroom walls, where AI is fast becoming part of everyday study.

Since 2022, generative AI tools such as ChatGPT or Grammarly have moved from the periphery to the mainstream of undergraduate study (Annamalai et al., 2023; Lam, 2024). On-demand correction, translation, conversational practice: these tools personalize instruction at a scale classroom teachers cannot easily reach (Baskara & Mukarto, 2023; Benek, 2025).

[1] *Corresponding author, Saigon Institute of Technology, HCMC, Vietnam; Email: trunght@saigontech.edu.vn*
[2] *ASEL Language School, Tay Ninh, Vietnam*
[3] *Ho Chi Minh City University of Technology, HCMC, Vietnam*
[4] *Associate Professor, Faculty of Foreign Languages, Nguyen Tat Thanh University, HCMC, Vietnam*

Consider the learner pursuing self-directed learning (SDL), "an awareness and control over ways to learn" (Curry et al., 2017, p. 18). For that learner, AI looks like the ideal extension of teacher-led instruction: scaffolding available anywhere, anytime, tailored to whatever question arises.

The empirical record on AI in EFL settings grows fast, but unevenly. Three limitations recur. First, most studies fix on a single tool, usually ChatGPT, and a single skill, usually writing (Pham & Cao, 2025). Second, many report what learners think AI is good for, not the strategic moves learners actually make once AI enters the picture (Wang et al., 2025; Yang & Li, 2024). Third, despite an obvious conceptual fit, vanishingly few anchor their analysis in established frameworks of language learning strategies or self-regulated learning (Zhang & Zou, 2024). The Vietnamese picture narrows further: classroom studies dominate, and gender's role in AI-mediated SDL remains largely unexamined (Ngo et al., 2025; Vo & Nguyen, 2024).

These gaps matter for two reasons, both tied to the conference theme. First, the EFL-to-ESL transition is, for the individual learner, a shift away from compliance with classroom routines toward autonomous, lifelong engagement with English; lacking evidence on how learners self-direct with AI, teacher educators cannot prepare future teachers to scaffold that shift responsibly. Equity is the second. Should male and female students draw on AI in systematically different ways, generic guidance will quietly leave some students under-served.

Accordingly, this study addresses two research questions:

RQ1. What AI-driven self-directed language learning strategies (AI-SDLLS) do Vietnamese EFL students use?

RQ2. What are gender differences in using AI-SDLLS in self-directed language learning?

To answer these questions, we adopt a convergent mixed-methods design at a vocational college in Ho Chi Minh City, pairing a questionnaire (N = 52) with semi-structured interviews (n = 16). Vocational-college learners merit particular attention here: relative to the four-year university students who dominate the Vietnamese evidence base, they more often combine study with paid work, enter with more uneven English proficiency, and depend on more variable digital access, conditions that plausibly shape both how readily they turn to AI and which strategies they can sustain. The study offers a localized portrait of AI-mediated SDL in a Vietnamese setting rarely visible in the international literature, alongside practical guidance for teachers, teacher educators, and policy actors readying students for English as a second language.

2. Literature Review

2.1. Self-Directed Learning and Language Learning Strategies

Self-directed learning is most often traced to Knowles (1975). He cast it as a process in which adults "take the initiative…in diagnosing their learning needs, formulating goals, identifying human and material resources, choosing and implementing appropriate learning strategies, and evaluating outcomes" (p. 18). Curry et al. (2017) later recast SDL for the language classroom as self-directed language learning (SDLL), foregrounding awareness, management, and autonomy. Pemberton and Cooker (2012) sharpened the picture: SDLL surfaces when "students take the initiative in setting their own goals, developing learning strategies, or employing various learning styles" (p. 203).

Strategic action sits at the core of SDLL. Building on O'Malley and Chamot (1990) and Oxford (1990, 2016), researchers typically sort language learning strategies into four families: cognitive (manipulating the language: translating, summarizing, repetition), metacognitive (planning, monitoring, and evaluating one's own learning), affective (managing emotions, motivation, and anxiety), and social (interacting and collaborating with others). Hawkins (2018) reworked this taxonomy into cognitive, affective, sociocultural-interactive, and meta-strategies. This study keeps the four-strategy framework. It dominates Vietnamese EFL research and yields categories that hold up across studies.

2.2. AI Tools in Language Learning

AI entered the SDLL toolkit precisely as learners began reaching for support beyond the classroom. Three classes of tools matter most here. Generative AI applications such as ChatGPT and Gemini generate human-like text that learners mine for explanation, paraphrasing, and feedback (Caruccio et al., 2024; Peres et al., 2023). Intelligent personal assistants like Alexa and Siri draw on natural language processing to "tailor the pace, approaches, content, and activities of learning to match the interests and needs of individual language learners" (Chen et al., 2021, p. 205). Then come mobile-assisted language learning (MALL) applications (Duolingo, Busuu, HelloTalk) which fold practice into everyday digital routines (Lai et al., 2022).

Empirical work links these tools to gains across the four skills. Writing benefits from AI support for idea generation, organization, grammar correction, and stylistic feedback (Abdalkader, 2022; Jensen et al., 2024; Zhang et al., 2023). In speaking, chatbots and speech-recognition systems open up low-stakes rehearsal and real-time pronunciation feedback (Rukiati et al., 2023). Listening widens too: AI-mediated platforms such as Netflix and YouTube (subtitles, translation, replay) expand exposure to authentic input (Suryana et al., 2020). For reading, dictionary-style lookups, summaries, and text-to-speech tools feed vocabulary growth and comprehension (Herawati et al., 2022). Yet most reviews caution that AI's pedagogical value rests less on the tool than on how learners use it (Mohebbi, 2024; Wang et al., 2025). This is where SDLL frameworks become useful.

2.3. AI-SDLL: A Strategic View of Learner Engagement

Framing AI use through this four-strategy lens clarifies what AI-SDLLS denotes: AI does not replace the strategy categories so much as extend each one. Cognitively, learners offload translation, grammar correction, summarizing, and vocabulary practice to AI, which returns immediate, low-cost feedback. Metacognitively, AI can scaffold goal-setting, planning, self-monitoring, and reflection, prompting learners to evaluate their answers or ask how to improve, though these uses demand more deliberate initiative. Affectively, AI offers low-stakes, non-judgmental practice that can ease anxiety and sustain motivation, while also risking over-reliance. Socially, AI serves as an interactional partner or a rehearsal space for human contact, from conversational chatbots to checking one's language before speaking with others. Read this way, AI-SDLLS are not new strategy types but familiar strategies re-mediated by AI affordances, which makes Oxford's (2016) taxonomy a productive frame for analyzing them.

Recent studies converge on three findings about AI-mediated SDL. First, learners reach for AI most readily on cognitive tasks (translation, correction, vocabulary lookup), where entry costs little and feedback arrives at once (Annamalai et al., 2023; Jeon, 2022). Second, metacognitive use is harder to cultivate without explicit instruction; left to their own devices, learners deploy AI reactively rather than for planning or monitoring (Wang et al., 2024; White, 2008). Third, affective benefits (lower anxiety, higher motivation) are real, yet ambivalent,

shadowed by recurrent risks of over-reliance and shallow engagement (Cotton et al., 2023; Özkan et al., 2024). Social use lags furthest behind. Conversational practice with chatbots stays thinly developed, reflecting both technological limits and learners' preference for human partners (Rukiati et al., 2023).

2.4. Gender and AI-Mediated Language Learning

Gender has long drawn interest in language learning strategy research. Female learners are sometimes reported as heavier users of metacognitive and social strategies (Alhaysony, 2017; Montero-SaizAja, 2021), though the findings waver across contexts. In AI-mediated studies the evidence thins considerably. Selim (2024) found that female Saudi undergraduates turned to Grammarly and GPT-3 to polish their academic writing more readily than male peers, even if the confidence gains stayed uncertain. Liu and Wang (2024) detected no significant gender difference in critical-thinking gains from AI in EFL literature classes. Reporting greater listening-comprehension difficulty among female learners, Huang (2024) tentatively linked it to less efficient processing strategies. Closer to home, Ngo et al. (2025) found no significant gender differences in overall AI use, yet pressed for finer-grained analysis of strategic patterns.

2.5. AI-SDLL in Vietnamese Higher Education

Vietnamese empirical work on AI-mediated SDL has expanded rapidly since 2023, but it clusters in two areas: students' perceptions of generative AI for writing tasks (Ho et al., 2024; Vo & Nguyen, 2024) and survey-based explorations of self-efficacy (Ngo et al., 2025). Such work establishes that Vietnamese undergraduates use AI often and regard it favorably. What it rarely does is examine the four-strategy structure typical of Oxford-style research, folding strategy use instead into broad categories. They also draw predominantly on four-year university samples; vocational colleges, which enroll a substantial share of post-secondary learners in Vietnam and differ in proficiency, time pressure, and digital infrastructure, are largely absent from the published record. The strategic profile of vocational EFL learners therefore remains an evidence gap with direct implications for an ESL-oriented system, where learners must self-regulate across longer time horizons and more diverse communicative settings.

2.6. The Present Study

This review exposes three connected gaps. The first is empirical: despite rapid growth in research on AI in Vietnamese EFL, no study has mapped the full four-strategy AI-SDLLS profile (cognitive, metacognitive, affective, and social) that learners enact when they self-direct with AI, as strategy use is typically folded into broad perception categories. The second is contextual: this evidence base draws overwhelmingly on four-year university samples, leaving vocational-college learners, who differ in proficiency, time pressure, and digital access, largely undocumented, even though they form a substantial share of Vietnam's post-secondary population. The third is methodological: gender and strategy patterns have rarely been examined with converging quantitative and qualitative evidence, so survey-level patterns remain unexplained by accounts of how learners actually use AI. The present study addresses these three gaps together and locates its findings within the conference theme of moving from EFL to ESL practice.

3. Methodology

3.1. Research Design

A convergent mixed-methods design (Creswell & Plano Clark, 2018) was used to address the two research questions. In a convergent design, quantitative and qualitative data are collected concurrently, analyzed separately, and then merged so that each strand can corroborate, explain, or complicate the other. This logic suited the study's dual aim: the closed-ended questionnaire was needed to establish the group-level frequency and patterning of AI-SDLLS, whereas the semi-structured interviews were needed to show how individual students actually enacted those strategies and reasoned about gender, detail that frequencies alone cannot convey. The two strands were given equal priority and were merged at the interpretation stage, where interview accounts were read against each subscale to judge whether they confirmed the survey pattern (for example, the prominence of cognitive uses), explained it (for example, why social uses lagged), or complicated it. Because the data came from a single institution, the design was intended to yield an in-depth, context-specific account rather than findings generalizable to all Vietnamese EFL undergraduates.

3.2. Research Context

The study was conducted at a vocational institution in Ho Chi Minh City. The college offers programs in English Language, Business Interpretation, Hospitality and Tourism Management, Logistics, and Information Technology. Communicative competence anchors the English Language program; instructors routinely recommend AI tools (Grammarly, QuillBot, ChatGPT, Grok, Copilot, Monica) as out-of-class support, and the college runs occasional workshops on AI use.

3.3. Participants

Participants were recruited through convenience sampling at a single vocational college; all were enrolled EFL students who volunteered to take part. The achieved sample of 52 is modest but adequate for an exploratory, descriptive study of strategy-use patterns; as noted below, however, it affords limited statistical power for detecting small gender differences. Fifty-two undergraduate EFL students, all aged 18 or older, completed the survey. Just over half (51.9%) fell into the 18–19 band; 26.9% were 20–21, and 21.2% were 22 or older. Males made up 53.8% and females 44.2%, with 1.9% identifying as another gender. Year level skewed early: 50.0% second-years, 38.5% first-years, and 11.5% in their third year or beyond. More than a third (38.5%) had studied English for five years or longer.

All respondents owned a personal learning device. ChatGPT was the most-used AI tool (88.5%), followed by Google Translate (59.6%), Duolingo (50.0%), and Grammarly (32.7%). Newer generative AI tools (Gemini, DeepSeek, Character.AI) were used by 3.8% or fewer. Writing was the most-supported skill (50.0% of participants used AI for it "often" or "always"), followed by vocabulary (46.2%), reading (46.1%), listening (40.3%), and speaking (34.6%). For the interview phase, 16 survey respondents were purposively selected from those who volunteered, balancing gender (seven male, nine female) and representing different year levels. Because these participants were volunteers, some volunteer bias, toward students more engaged with AI or more confident in English, cannot be ruled out and is acknowledged as a limitation. These profile characteristics (year level, length of English study, near-universal device ownership, the concentration of use in a few tools, notably ChatGPT, and the skew toward writing and vocabulary support) are not merely descriptive; each plausibly conditions which AI-SDLLS students adopt and is therefore treated as interpretive context for the results rather than as a controlled variable.

3.4. Instruments

3.4.1. Questionnaire

A 43-item questionnaire was developed for this study, organized into five sections: demographics, AI tool usage, and four AI-SDLLS subscales (cognitive: 9 items; metacognitive: 7 items; affective: 6 items; social: 7 items), each rated on a 5-point Likert scale (1 = Never, 5 = Always). The subscales were constructed by taking Oxford's (2016) four strategy families and Knowles' (1975) SDL principles as the conceptual scaffold and writing items that re-expressed each strategy as a concrete AI-mediated behavior. Cognitive items captured using AI for translation, grammar checking, summarizing, and vocabulary practice; metacognitive items captured AI-supported goal-setting, planning, monitoring, and reflection; affective items captured using AI to manage motivation, enjoyment, and anxiety; and social items captured using AI to prepare for or mediate interaction with others. Items were newly written for this study, with wording informed by behaviors reported in prior AI-in-EFL research rather than adopted from a single validated scale, as no existing instrument operationalized AI-SDLLS across all four families.

Content validity was established through expert review: two EFL instructors independently judged whether each item belonged to its intended strategy family and was clearly worded, and items were revised accordingly. The instrument was then pilot-tested with three students to confirm clarity. Internal consistency was strong across all four subscales (see Table 1). Construct validity was not formally tested through factor analysis, because the sample was too small to support a stable factor solution; this is acknowledged as a limitation of instrument validation.

Table 1

Internal Consistency of the AI-SDLLS Subscales

Subscale	k	Cronbach's α
Cognitive	9	.85
Metacognitive	7	.84
Affective	6	.84
Social	7	.85

Note. k = number of items; N = 52.

3.4.2. Semi-structured Interview

A 20-question protocol covered (a) introduction, (b) AI tool usage, and (c) AI-SDLLS practice across the four strategy families. Two sources shaped the protocol: Benson's (2013) autonomy framework and recent empirical work on AI in EFL contexts (Annamalai et al., 2023). Optional follow-up prompts nudged participants toward concrete examples. Before administration, two students piloted the protocol and two further instructors reviewed it. The full interview protocol appears in Appendix A.

3.5. Data Collection and Analysis

Collection ran two weeks through Google Forms. Survey respondents completed the instrument at their own pace; interview participants, having signed a consent statement, received the form and replied asynchronously in writing. This format ruled out real-time probing but permitted thoughtful and reflective responses.

Quantitative data were analyzed in IBM SPSS Statistics 31. Descriptive statistics (M, SD) were computed per item and per subscale. Internal consistency was tested with Cronbach's α. Independent-samples t-tests compared male and female participants on the four subscale means (the one participant identifying as another gender was excluded from this comparison, leaving n = 51). Levene's test indicated equal variances for all four subscales, so equal-variance results are reported, and Cohen's d indexed effect size. Given the modest sample, these tests had limited power to detect small effects, so non-significant results are read as consistent with either genuine similarity or insufficient power, not as evidence of no difference.

Qualitative data were analyzed using Braun and Clarke's (2006) six-phase thematic analysis: familiarization, initial coding, theme generation, theme review, definition, and reporting. Codes were generated both inductively and deductively, organized around the four-strategy framework. Trustworthiness was supported through peer debriefing, an audit trail, and triangulation with the quantitative findings. Formal member checking was not undertaken; because interviews were administered in writing through Google Forms, participants' original responses were retained verbatim and consulted directly during analysis, and excerpts are reported with minimal editing to preserve participants' wording. In peer debriefing, research-team members independently reviewed a subset of coded responses and discussed discrepancies until the coding scheme and theme definitions were agreed, which sharpened several category boundaries (for example, separating metacognitive monitoring from cognitive checking). Themes were merged with the quantitative results during interpretation. Participants are identified anonymously as S1–S16.

4. Results

4.1. AI-SDLLS Use Across the Four Strategy Families

Across the sample, students reported using AI-SDLLS at moderate levels overall. Cognitive strategies (M = 3.38, SD = .80) were the most frequently used. Metacognitive (M = 3.09, SD = .85) and affective strategies (M = 3.09, SD = .89) followed at very similar levels, and social strategies (M = 3.02, SD = .89) ranked lowest. The clearest feature of these data is the gap between cognitive use and the other three families; the differences among metacognitive, affective, and social use, by contrast, are small (no greater than 0.07 on the 5-point scale). Because no inferential test was conducted across the four subscale means, this ordering is reported as a descriptive tendency rather than a statistically established hierarchy (see Table 2).

Table 2

EFL Students' Use of AI-SDLLS

No.	Components	N = 52	
		M	**SD**
1	Cognitive	3.38	.80
2	Metacognitive	3.09	.85
3	Affective	3.09	.89
4	Social	3.02	.89
	Average	**3.15**	**.86**

*Note. Scale: 1 = Never, 5 = Always.

4.2. Cognitive Strategies

Within the cognitive subscale (Table 3), translating between English and Vietnamese (Cog2: M = 3.75, SD = 1.12) and grammar checking (Cog8: M = 3.75, SD = 1.20) were the most frequent uses. Vocabulary practice (Cog1: M = 3.58, SD = 1.09) and exploring alternative expressions (Cog6: M = 3.58, SD = 1.05) were also common. Less frequent were sentence-construction practice with new words (Cog5: M = 3.00, SD = 1.25) and pronunciation practice via text-to-speech (Cog7: M = 2.83, SD = 1.32).

Table 3

EFL Students' Use of Cognitive Strategies

No.	Items	N = 52	
		M	SD
Cog1	I use AI to practice new vocabulary.	3.58	1.09
Cog2	I use AI to translate between English and Vietnamese.	3.75	1.12
Cog3	I use AI to summarize English texts to understand better.	3.39	1.16
Cog4	I use AI to make practice questions for myself.	3.19	1.14
Cog5	I use AI to make sentences with new vocabulary items.	3.00	1.25
Cog6	I use AI to know different ways to express my ideas.	3.58	1.05
Cog7	I listen to AI text-to-speech to listen to and copy correct pronunciation.	2.83	1.32
Cog8	I use AI to check my grammar when I write.	3.75	1.20
Cog9	I use AI tools to compare my writing with better examples.	3.35	1.20

Interview accounts supported the quantitative pattern. S1 explained, "To be honest, I am really bad at writing and grammar, so I often use AI tools to practice them." S3 added, "I often use AI tools to practice reading. It helps me read faster, understand new words, and check whether I did it right or wrong." These accounts illustrate why cognitive support emerged as AI's most prominent role for participants: it addressed perceived weaknesses with immediate, low-effort feedback.

4.3. Metacognitive Strategies

For metacognitive strategies (Table 4), one item alone crossed the "often" threshold: evaluating answers after exercises (MetaCog6: M = 3.46, SD = 1.15). Reviewing mistakes to learn from them followed it narrowly (MetaCog3: M = 3.44, SD = 1.13). Long-term planning fared worse. Arranging a learning schedule through AI (MetaCog2: M = 2.77, SD = 1.23) and soliciting AI suggestions for further improvement (MetaCog7: M = 2.79, SD = 1.19) drew the lowest ratings.

Table 4

EFL Students' Use of Metacognitive Strategies

No.	Items	N = 52	
		M	SD
MetaCog1	I use AI to set learning goals.	2.92	1.27

MetaCog2	I plan my study by asking AI to make a study schedule for me.	2.77	1.23
MetaCog3	I use AI tools to review my mistakes and learn from them.	3.44	1.13
MetaCog4	I use AI tools to choose exercises that match my level.	3.15	1.27
MetaCog5	I ask AI for feedback on my speaking or writing.	3.12	1.15
MetaCog6	I use AI to evaluate my answers after finishing exercises.	3.46	1.15
MetaCog7	I reflect on my study by asking AI, "How can I do better next time?"	2.79	1.19

The interviews exposed the same lopsidedness. S2 used AI mainly to diagnose weaknesses: "I'll ask for a sample test and AI will check it and explain what my problems are and what I need to improve." A minority did set goals or monitor progress with AI. S10 reported, "I set [a goal] to reach 70+ at PTE in 6 months, and like a coach every day I [log] the day I learn," and S9 explained, "When I write a paragraph, I ask ChatGPT to give me feedback and show my mistakes. It tells me what parts are weak, like grammar or vocabulary." Yet sustained planning remained the exception; as one student admitted, "I make plans for using AI, but I cannot always follow them." Diagnosis and on-the-spot evaluation came easily; disciplined, forward-looking planning did not.

4.4. Affective Strategies

Affective strategy use (Table 5) was moderate. Three uses clustered closely: AI sustained motivation through immediate feedback (Affect4: M = 3.27, SD = 1.14), made learning enjoyable via gamified applications (Affect3: M = 3.17, SD = 1.26), and offered encouragement through achievements and rewards (Affect5: M = 3.17, SD = 1.09). Reducing the fear of mistakes rated slightly higher (Affect6: M = 3.33, SD = 1.13). Rarest of all was building speaking confidence through chatbot interaction (Affect1: M = 2.67, SD = 1.28).

Table 5

EFL Students' Use of Affective Strategies

No.	Items	N = 52	
		M	SD
Affect1	I talk with AI chatbots to build confidence in speaking.	2.67	1.28
Affect2	I use AI tools to reduce stress before tests.	2.94	1.26
Affect3	I use AI games or apps to make learning fun.	3.17	1.26
Affect4	I use AI tools to stay motivated by seeing quick results.	3.27	1.14
Affect5	Learning with AI helps me stay motivated with rewards for reaching milestones.	3.17	1.09
Affect6	Learning with AI helps me feel less afraid of making mistakes.	3.33	1.13

Interview responses split along the same lines. For some, AI relieved anxiety: S13 valued that "it's just an AI, not a real human, so I'm not afraid to be judged if I have a mistake," and S12 said, "I feel relaxed and motivated because it helps me with everything, from checking mistakes to giving me new exercises." Others were cooler or warier. S10 reported feeling "nothing... [I treat] it as just a tool, nothing more, nothing less," while S2 felt, "Sometimes I feel really bored because myself depend on AI." Such ambivalence hints that AI can ease anxiety yet, overused, may dull intrinsic interest.

4.5. Social Strategies

Social strategies (Table 6) were the least frequently used. Students did use AI to generate ideas for group activities (Social4: M = 3.48, SD = 1.02), to check English before social interaction (Social5: M = 3.40, SD = 1.16), and to clarify peers' or teachers' English (Social7: M = 3.12, SD = 1.34). However, treating AI as a speaking partner (Social1: M = 2.54, SD = 1.31) and practicing virtual conversations (Social3: M = 2.60, SD = 1.23) were rare.

Table 6

EFL Students' Use of Social Strategies

No.	Items	N = 52	
		M	SD
Social1	I use AI as a speaking partner.	2.54	1.31
Social2	I share AI-generated ideas with my friends to start conversations.	2.87	1.31
Social3	I use AI tools as virtual friends to practice real-life conversation.	2.60	1.23
Social4	I use AI tools for ideas to complete group activities.	3.48	1.02
Social5	I use AI to check my English before talking with others.	3.40	1.16
Social6	I ask AI to suggest topics to discuss with other people.	3.17	1.18
Social7	I use AI to help me understand classmates' or teachers' English.	3.12	1.34

The interviews echoed the numbers. Where students used AI socially, it was mostly to prepare for human interaction: S13 explained, "I often use AI to prepare for discussing with my classmates or answering questions from teachers; I ask AI for the main points, then I widen ideas from them." A few embraced it as a conversational partner. S14 said, "When it comes to speaking, I tell ChatGPT to make a conversation for me and we talk like friends," and S8 noted, "I think we can talk to AI anytime, anywhere." Even so, most ultimately favored peers and teachers for genuine social practice; as S3 put it, "a good partner for speaking practice is a classmate, because real people will be better than AI."

4.6. Gender Differences in AI-SDLLS

Independent-samples t-tests revealed no statistically significant gender differences on any of the four subscales (Table 7). Levene's test indicated equal variances throughout, so equal-variance results are reported, t(49). Mean differences were small, no larger than about 0.15 on the 5-point scale, and effect sizes were negligible to small (d = -.18 to .13). Female participants edged ahead on cognitive (M = 3.44 vs. 3.30) and affective (M = 3.16 vs. 3.01) use, and male participants on metacognitive use (M = 3.13 vs. 3.01), while social-strategy use barely differed (M = 3.05 vs. 3.00). These null results should be read with caution: at this sample size the t-tests had limited power to detect small effects, so the absence of significant differences is consistent with either genuine similarity or insufficient power. The interviews likewise surfaced no marked gendered patterns, with male and female participants alike emphasizing cognitive uses; however, because the interviews were not designed primarily to elicit gendered accounts, this convergence is treated as broadly consistent with the survey rather than as independent confirmation of no gender difference.

Table 7

Differences in EFL Students' Use of AI-SDLLS by Gender

Subscale	Male M (SD)	Female M (SD)	*t* (df)	*p*	*d*
Cognitive	3.30 (0.85)	3.44 (0.75)	-0.63(49)	.53	-0.18
Metacognitive	3.13 (0.93)	3.01 (0.76)	0.48(49)	.64	0.13
Affective	3.01 (1.02)	3.16 (0.72)	-0.61(49)	.55	-0.17
Social	3.00 (0.98)	3.05 (0.81)	-0.20(49)	.85	-0.05

Note. N = 51 (male = 28, female = 23); the one participant identifying as another gender was excluded from this comparison. Equal variances assumed (Levene's $p > .05$ for all subscales). d = Cohen's d.

5. Discussion

This study set out to map how Vietnamese EFL undergraduates draw on AI-SDLLS, and whether gender shapes that use. Three patterns stand out, each worth reading against Vietnam's broader move toward ESL.

5.1. A Strategy Hierarchy Anchored in Cognitive Use

Cognitive strategies dominated. This replicates earlier findings that learners first exploit AI for concrete, surface-level scaffolding like grammar correction, translation, idea generation (Annamalai et al., 2023; Jensen et al., 2024). For Oxford (2016), such strategies form the most accessible entry point into strategic learning, a reading that fits our participants' A2–B1 proficiency. Yet an ESL-oriented system carries a particular risk: learners stall at this cognitive layer. AI becomes a pair of "task hands" rather than a partner in extended, autonomous learning.

5.2. Underdeveloped Metacognitive and Affective Use

Metacognitive use was uneven. Students turned to AI to evaluate and review finished work, yet rarely to plan ahead or set long-term goals. White (2008) and Lai et al. (2022) noted the same: metacognitive growth seldom happens without explicit support. Vietnam's traditional teacher-centered classroom may sharpen the problem. Where instruction prizes accuracy and teacher direction over self-regulation, students reach AI-mediated SDL with few planning habits in place (Ngo et al., 2025).

Affective use proved equally ambivalent. For some students, AI worked as a non-judgmental partner that took the edge off anxiety (Lai et al., 2022). Others described boredom and dependence, an echo of Cotton et al. (2023) and Amaro et al. (2023) on shallow engagement and over-reliance. Juggling jobs alongside coursework, with little time for deliberate practice, the vocational students here look especially prone to passive dependence on AI.

5.3. The Social Gap

Social strategies lagged furthest behind. Few students used AI as a conversational partner, and many called the chatbot exchanges "robotic," squaring with Rukiati et al. (2023) on how poorly chatbots reproduce human affective nuance. Two contextual factors likely contribute. Free-tier tools seldom offer high-quality voice interaction, leaving cost-conscious learners with text alone. Low oral proficiency, in turn, makes sustained AI dialogue hard to keep going. Where authentic spoken interaction sits at the center of an ESL-oriented system, this gap is the most pedagogically urgent.

5.4. Gender: Small and Context-Dependent

No statistically significant gender differences emerged. This contrasts with earlier strategy research (e.g., Alhaysony, 2017), though it aligns with more recent Vietnamese work (Ngo et al., 2025). Two cautions frame this null result. First, the modest sample gave the t-tests limited power to detect small effects, so the finding is best read as an absence of evidence for sizable gender differences rather than evidence of their absence. Second, the interviews were not designed primarily to elicit gendered accounts, and they too surfaced no marked pattern. Taken together, these results suggest that in a relatively homogeneous vocational sample, gender is a weaker predictor of AI-SDLLS than factors such as proficiency, digital literacy, time pressure, and access to tools, variables that plausibly shape strategy use more directly and that cut across gender. This does not render gender irrelevant, since small directional differences could matter at scale; rather, it implies that AI-strategy instruction should be designed for all students while attending to the conditions that more clearly differentiate them.

5.5. Implications for the EFL-to-ESL Transition

Read together, these patterns of AI use, heavily weighted toward task-level cognitive support and lighter on planning, self-regulation, and interaction, are consistent with, though not direct proof of, a broadly task-oriented use of AI rather than the autonomous, communicative use implied by an ESL orientation. It bears emphasizing that the study measured how students use AI strategically, not their broader stance toward English or their identity as EFL or ESL learners; the link to the EFL-to-ESL transition is therefore interpretive rather than demonstrated. With that caveat, the findings suggest that supporting an ESL-oriented shift will require deliberate scaffolding of metacognitive and social strategies, not merely access to more powerful tools.

5.6. Pedagogical Moves for Teachers and Programs

Three concrete moves emerge from the findings. The first is targeted strategy instruction. Because cognitive use of AI is already established, classroom time can be devoted to the layers that are underdeveloped: planning, monitoring, and reflective evaluation of AI feedback. Brief in-class routines (i.e. setting a weekly AI-use goal, logging the prompts that worked, comparing AI suggestions with peer feedback) convert AI from a hidden out-of-class crutch into a visible object of metacognitive learning. These routines align with the autonomy framework of Benson (2013) and with recent work on AI-supported self-regulation (Mohebbi, 2024).

The second is the deliberate design of social practice. Where chatbot conversations feel robotic to learners, teachers can scaffold hybrid activities: peer dialogues informed by AI preparation, role-plays that mix human and AI interlocutors, and small-group reviews of AI-generated transcripts. Such activities preserve the human relational dimension that learners value while extracting the linguistic affordances of AI. They are also well aligned with strand 6 of the conference theme, which positions technology and AI as supports for ESL development rather than as ends in themselves.

The third is teacher preparation. Because students already access AI for cognitive support, teacher development should target not basic access but higher-order use: designing reflective prompts, evaluating the accuracy and usefulness of AI feedback, comparing AI output with peer and teacher feedback, using AI ethically, and recognizing and countering overreliance. Pre-service and in-service teachers need experiential training that equips them to model this kind of AI literacy rather than merely police AI use; teacher-education programs can fold in brief modules on prompt design, ethical evaluation of AI outputs, and strategy-

explicit assignment design (Eldin, 2024). Skip this preparation, and the gap between policy aspirations for ESL competence and classroom realities will only widen.

These moves must be set against a backdrop of unequal access. Although every student in this sample owned a learning device, free-tier tools, premium features such as advanced voice interaction, and, critically for working vocational students, stable time to practice are not evenly distributed. Equitable AI integration therefore means more than supplying tools; it means attending to which learners can realistically reach the higher-quality, interactive features that an ESL-oriented use of AI depends on.

6. Conclusion

This mixed-methods study examined how Vietnamese EFL undergraduates at a vocational college in Ho Chi Minh City use AI-driven self-directed language learning strategies, and whether that use varies by gender. A consistent hierarchy emerged: cognitive strategies dominate, metacognitive and affective strategies see moderate use, and social strategies remain underdeveloped. Independent-samples t-tests turned up no significant gender differences across the four subscales.

Three pedagogical implications follow. First, classroom programs should treat AI literacy as an explicit learning objective, set beside language skills. Students need guided practice in moving past surface-level correction toward planning, monitoring, and critically weighing AI feedback, the very activities at the heart of self-regulated learning. Second, teachers need professional development that frames AI as a pedagogical resource rather than a competitor or threat; absent such support, principled classroom integration stays out of reach. Third, equity deserves explicit attention. Language proficiency, time pressure, and uneven access to premium tools shape who benefits most from AI, and the stakes run highest in vocational and under-resourced settings. Unless these conditions are addressed, Vietnam's move toward English as a second language risks becoming stratified by digital access rather than inclusive.

Several limitations warrant acknowledgment. Modest in size and drawn through convenience and voluntary sampling from a single vocational college, the sample constrains generalizability to other Vietnamese institutional contexts, and its size gave the gender comparison limited statistical power to detect small effects. The data were entirely self-reported, with no direct evidence from AI-interaction logs, learning diaries, or language-performance measures, and the asynchronous interview format limited real-time probing of nuanced responses. Future research could triangulate self-reports with proficiency measures, AI-interaction logs, and longitudinal designs tracking how AI-SDLLS evolve over a program. Comparative work across institutional types (secondary schools, universities, and vocational colleges) would further clarify how contextual factors shape AI-SDLLS adoption as Vietnam continues its transition from EFL to ESL practice.

Three directions for future research deserve specific attention. The first calls for intervention studies, testing whether short, targeted instruction in metacognitive and social AI strategies measurably shifts the strategy profile reported here. The second requires larger, stratified samples to see whether the small directional gender differences here hold up as statistically reliable in more powered designs. The third points to qualitative work tracking individual learners across a semester, complementing cross-sectional findings by showing how AI-SDLLS evolve in response to course demands, assessment formats, and teacher intervention.

Despite these limitations, the study contributes a localized portrait of AI-mediated SDL in a context rarely visible in the international literature. It speaks directly to the HUFLIT INTESOL 2026 theme, showing what the EFL-to-ESL transition looks like at the level of individual learners working with AI tools, and what teacher educators, curriculum designers, and policy actors can do to support it.

References

Abdalkader, S. M. A. (2022). Using artificial intelligence to improve writing fluency for the preparatory stage students in distinguished governmental language schools. *Egyptian Journal of Educational Sciences, 2*(2), 39–70. https://web.archive.org/web/20250805013449/https://ejes.journals.ekb.eg/article_270694.html

Alhaysony, M. (2017). Language learning strategies use by Saudi EFL students: The effect of duration of English language study and gender. *Theory and Practice in Language Studies, 7*(1), 18–28. https://doi.org/10.17507/TPLS.0701.03

Amaro, I., Della Greca, A., Francese, R., Tortora, G., & Tucci, C. (2023). AI unreliable answers: A case study on ChatGPT. In H. Degen & S. Ntoa (Eds.), *Lecture notes in computer science* (pp. 23–40). Springer. https://doi.org/10.1007/978-3-031-35894-4_2

Annamalai, N., Rashid, R. A., Hashmi, U. M., Mohamed, M., Alqaryouti, M. H., & Sadeq, A. E. (2023). Using chatbots for English language learning in higher education. *Computers and Education: Artificial Intelligence, 5*, 100153. https://doi.org/10.1016/j.caeai.2023.100153

Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics, 7*(2), 343–358. https://doi.org/10.21093/ijeltal.v7i2.1387

Benek, K. (2025). EFL learners' and teachers' perceptions of AI-powered language learning technologies: Benefits and challenges. *International Journal of Instruction, 18*(2), 103–120. https://doi.org/10.29333/iji.2025.1827a

Benson, P. (2013). *Teaching and researching: Autonomy in language learning* (2nd ed.). Routledge. https://doi.org/10.4324/9781315833767

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Caruccio, L., Cirillo, S., Polese, G., Solimando, G., Sundaramurthy, S., & Tortora, G. (2024). Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot. *Expert Systems with Applications, 235*, 121186. https://doi.org/10.1016/j.eswa.2023.121186

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modeling and knowledge mapping. *Educational Technology & Society, 24*(1), 205–222. https://www.jstor.org/stable/26977868

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Curry, N., Mynard, J., Noguchi, J., & Watkins, S. (2017). Evaluating a self-directed language learning course in a Japanese university. *International Journal of Self-Directed Learning, 14*(1), 17–36. http://docs.wixstatic.com/ugd/dfdeaf_385a2e4d19254f968487b6058464e00c.pdf

Eldin, A. H. S. (2024). Using artificial intelligence in EFL teacher education programs. *Sustainability Education Globe, 2*(1), 26–35. https://web.archive.org/web/20250315173243/https://seg.journals.ekb.eg/article_348587.html

Hawkins, M. W. (2018). Self-directed learning as related to learning strategies, self-regulation, and autonomy in an English language program: A local application with global implications. *Studies in Second Language Learning and Teaching, 8*(2), 445–469. https://doi.org/10.14746/ssllt.2018.8.2.12

Herawati, D. R. N., Widajati, W., & Sartinah, E. P. (2022). The role of text-to-speech assistive technology to improve reading ability in e-learning for ADHD students. *Journal of ICSAR, 6*(2), 169–174. https://doi.org/10.17977/um005v6i22022p169

Ho, H. B. N., Ho, H. B. N., & Thai, C. D. (2024). EFL students' perceptions and practices of using ChatGPT for developing English argumentative essay writing skills. *European Journal of Alternative Education Studies, 9*(1). https://doi.org/10.46827/ejae.v9i1.5341

Huang, S.-H. (2024). Exploring gender and language proficiency variations in English listening comprehension difficulties (LCDs) among college EFL (English as a foreign language) learners. *Language Education & Assessment, 7*(1), 1243. https://doi.org/10.29140/lea.v7n1.1243

Jensen, L. X., Buhl, A., Sharma, A., & Bearman, M. (2024). Generative AI and higher education: A review of claims from the first months of ChatGPT. *Higher Education*, *89*(4)1145-1161. https://doi.org/10.1007/s10734-024-01265-3

Jeon, J. (2022). Exploring a self-directed interactive app for informal EFL learning: A self-determination theory perspective. *Education and Information Technologies, 27*(4), 5767–5787. https://doi.org/10.1007/s10639-021-10839-y

Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Association Press.

Lai, Y., Saab, N., & Admiraal, W. (2022). Learning strategies in self-directed language learning using mobile technology in higher education: A systematic scoping review. *Education and Information Technologies, 27*(6), 7749–7780. https://doi.org/10.1007/s10639-022-10945-5

Lam, T. Q. A. (2024). AI chatbots in English language learning: A critical review. *Journal of Knowledge Learning and Science Technology, 3*(2), 185–195. https://doi.org/10.60087/jklst.vol3.n2.p195

Liu, W., & Wang, Y. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*. https://doi.org/10.1111/ejed.12804

Mohebbi, A. (2024). Enabling learner independence and self-regulation in language education using AI tools: A systematic review. *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2024.2433814

Montero-SaizAja, A. (2021). Gender-based differences in EFL learners' language learning strategies and productive vocabulary. *Theory and Practice of Second Language Acquisition, 7*(2), 83–107. https://doi.org/10.31261/TAPSLA.8594

Ngo, C.-L., Bui, N. V. C., Nguyen, P. N. N., & Nguyen, T. M. P. (2025). Exploring Vietnamese EFL learners' generative AI usage patterns, self-efficacy and self-directed learning: Preliminary findings. *Proceedings of the 13th OpenTESOL 2025 International Hybrid Conference*, Ho Chi Minh City Open University. https://www.researchgate.net/publication/396498666_Exploring_Vietnamese_EFL_learners'_generative_ai_usage_patterns_self-efficacy_and_self-directed_learning_preliminary_findings

Nguyen, P. B. T., & Tran, N. B. C. (2025). A SWOT analysis of artificial intelligence in English language teaching and learning in Southern Vietnam. In S. Mistretta (Ed.), *Education and human development*. IntechOpen. https://doi.org/10.5772/intechopen.1012477

O'Malley, J. M., & Chamot, A. U. (2012). *Learning strategies in second language acquisition*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524490

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle.

Oxford, R. L. (2016). *Teaching and researching language learning strategies: Self-regulation in context*. Routledge. https://doi.org/10.4324/9781315719146

Özkan, E. K., Erdemir, N., & Coşkun, D. (2024). A systematic review of EFL teachers' perspectives on artificial intelligence technologies. *Ihlara Eğitim Araştırmaları Dergisi, 9*(2), 150–168. https://doi.org/10.47479/ihead.1535035

Pemberton, R., & Cooker, L. (2012). Self-directed learning: Concepts, practice, and a novel research methodology. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning* (pp. 203–219). Palgrave Macmillan. https://doi.org/10.1057/9781137032829_14

Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing, 40*(2), 269–275. https://doi.org/10.1016/j.ijresmar.2023.03.001

Pham, M. T., & Cao, T. X. T. (2025). The practice of ChatGPT in English teaching and learning in Vietnam: A systematic review. *International Journal of TESOL & Education, 5*(1), 50–70. https://doi.org/10.54855/ijte.25513

Rukiati, E., Wicaksono, J. A., Taufan, G. T., & Suharsono, D. D. (2023). AI on learning English: Application, benefit, and threat. *Journal of Language, Communication and Tourism, 1*(2), 32–40. https://doi.org/10.25047/jlct.v1i2.3967

Selim, A. S. M. (2024). The transformative impact of AI-powered tools on academic writing: Perspectives of EFL university students. *International Journal of English Linguistics, 14*(1), 14. https://doi.org/10.5539/ijel.v14n1p14

Suryana, I., Asrianto, A., & Murwantono, D. (2020). Artificial intelligence to master English listening skills for non-English major students. *Journal of Languages and Language Teaching, 8*(1), 48–59. https://doi.org/10.33394/jollt.v8i1.2221

Vo, T. K. A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/fr1rkj58

Wang, C., Li, Z., & Bonk, C. (2024). Understanding self-directed learning in AI-assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence, 6*, 100247. https://doi.org/10.1016/j.caeai.2024.100247

Wang, Y., Zhang, T., Yao, L., & Seedhouse, P. (2025). A scoping review of empirical studies on generative artificial intelligence in language education. *Innovation in Language Learning and Teaching*, 1–28. https://doi.org/10.1080/17501229.2025.2509759

White, C. (2008). Language learning strategies in independent language learning: An overview. In T. W. Lewis & M. S. Hurd (Eds.), *Language learning strategies in independent settings* (pp. 3–24). Multilingual Matters.

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System, 124*, 103351. https://doi.org/10.1016/j.system.2024.103351

Zhang, R., & Zou, D. (2022). Self-regulated second language learning: A review of types and benefits of strategies, modes of teacher support, and pedagogical implications. *Computer Assisted Language Learning, 37*(4), 720–765. https://doi.org/10.1080/09588221.2022.2055081

Zhang, R., Zou, D., & Cheng, G. (2023). Chatbot-based training on logical fallacy in EFL argumentative writing. *Innovation in Language Learning and Teaching, 17*(5), 932–945. https://doi.org/10.1080/17501229.2023.2197417

Bionote

Huynh Thanh Trung is an English lecturer at the Saigon Institute of Technology (SaigonTech) and an M.A. student in English Language Studies at HUTECH University. Moving from engineering to linguistics and literature, his research explores the synergy between technology and pedagogy. His academic interests include AI-driven self-directed learning, digital literacy, and literature education. He focuses on developing frameworks that help students utilize emerging technologies, fostering independence in both practical language acquisition and the study of literary works.

Chau Thi Kim Ngan is an English teacher at ASEL Language School and an MA student English Language at HUTECH University. Her primary research interests encompass English language pedagogy, AI-driven self-directed learning, and the innovative integration of AI tools in language education. Her recent work focuses on utilizing AI in teaching methods to enhance EFL learners' speaking performance and learning attitudes.

Vo Thi Tien is an M.A. student in English language studies at HUTECH University. Her research interests regarding AI advancements in language education, AI-driven self-directed learning, and TESOL methodology. She would like to develop students' digital literacy, which promotes their self-study motivation and enhances language competence.

Tran Quoc Thao is working as a lecturer of English linguistics. His research areas are English language studies, TESOL methodology, and teacher professional development. He has served as a keynote speaker and an invited reviewer.

Appendix

Semi-Structured Interview Protocol

Part 1. Background. 1. Can you tell me a little about yourself, your name and what you study at Saigon College? (Why did you choose to study English?) 2. Do you enjoy learning English? Why or why not? (What part of English learning do you like best, e.g., speaking, writing?)

Part 2. AI Tool Use. 3. Which AI tools (e.g., Duolingo, ChatGPT, Grammarly, QuillBot, ELSA) do you use for learning English? (Why these? Are there others you tried but did not like?) 4. How often do you use AI tools for English learning, and when (e.g., at home, after class)? (Can you give an example?) 5. What English skills do you practice with AI tools (e.g., listening, speaking, reading, writing)? (Which skill the most? Why?)

Part 3. Cognitive Strategies. 6. Do you use AI tools to learn new words or grammar? How? 7. Can you give an example of when AI helped you understand a text or a difficult sentence? 8. How do you use AI tools to practice writing or speaking better? 9. Have you ever used AI to make practice exercises for yourself? What kind?

Part 4. Metacognitive Strategies. 10. Do you use AI to plan your study (e.g., daily goals, schedules)? Can you describe how? 11. How do you use AI to check your progress or find out your weaknesses? 12. Do you ever ask AI for advice about how to improve your English? What answers did it give? 13. How do you use AI to look back at your mistakes and learn from them?

Part 5. Affective Strategies. 14. How do you feel when using AI tools to learn English (happy, relaxed, stressed, motivated)? 15. Do you feel more confident when you use AI to practice speaking or writing? Why or why not? 16. Do you use AI tools for fun learning (games, songs, videos)? Which ones?

Part 6. Social Strategies. 17. Have you ever used AI to practice conversations in English? What was it like? 18. Do you use AI to prepare for talking with classmates or teachers? How? 19. Have you ever used AI to help you in group work or discussions? Can you share an example? 20. Do you think AI can be a good partner for speaking practice? Why or why not?

#180 AI-SUPPORTED PROBLEM-BASED LEARNING FOR ENHANCING UNIVERSITY STUDENTS' PUBLIC SPEAKING PERFORMANCE

Thi Thao Trang Nguyen[1]

Abstract

Driven by the shift towards learner-centered education and knowledge construction through solving real-world problems, Problem-Based learning (PBL) has demonstrated its effectiveness globally over the past five decades. This study investigates the integration of AI-driven virtual simulations and automated feedback into a Learning Management System (LMS) within a Problem-Based Learning (PBL) framework to improve students' public speaking skills. Employing a quasi-experimental design, the research involved two intact classes at Ho Chi Minh City University of Technology, divided into an experimental group (AI-supported LMS instruction) and a control group (traditional instruction). Data was gathered through pre-and-post performance assessments and an online survey evaluating students' perceptions, engagement levels, and attitudes toward AI-integrated learning. The results indicate that embedding AI tools within the PBL cycle significantly improved students' public speaking performance in fluency and confidence, while fostering greater learner autonomy.

Keywords: Problem-Based Learning (PBL), artificial intelligence, LMS, innovation, public speaking.

1. Introduction

In an era where communication competence is increasingly recognized as a critical graduate attribute, the ability to speak confidently and effectively in public has become essential for university students. However, despite its importance, public speaking remains a persistent challenge for many learners, often constrained by anxiety, limited practice opportunities, and insufficient individualized feedback. Traditional instructional approaches, while structured, frequently fail to provide the dynamic and learner-centered environment necessary for meaningful skill development. Against this backdrop, Problem-Based Learning (PBL) has gained prominence as a pedagogical approach that situates learning within authentic problem-solving contexts, encouraging active engagement, critical thinking, and self-directed learning. At the same time, the rapid advancement of Artificial Intelligence (AI) is reshaping educational landscapes by enabling adaptive learning, real-time feedback, and personalized support on a scale. The convergence of these two approaches presents a compelling opportunity to reimagine how complex skills such as public speaking are taught and learned. This study explores the integration of AI-supported tools within a PBL framework throughout Learning Management System (LMS) to enhance university students' public speaking performance. Students can practice speaking skills whenever they have a free time at home. By embedding intelligent feedback mechanisms and interactive simulations into a structured problem-solving cycle, the proposed approach aims to create a responsive and supportive learning environment where students can practice, reflect, and improve iteratively. The research seeks to contribute to emerging discussions on technology-enhanced pedagogy while

[1]*Foreign Language Center – HCMC University of Technology, Email: trangnguyen@hcmut.edu.vn*

addressing a critical gap in developing both competence and confidence in public speaking among higher education learners.

To achieve these objectives, the study addresses the following research question:

What is the impact of an AI-supported Learning Management System (LMS) integrated within a Problem-Based Learning (PBL) framework on university students' public speaking performance, confidence, motivation, and learner autonomy?

2. Literature Review

2.1. Theoretical Foundations of Problem-Based Learning (PBL)

Problem-Based Learning (PBL) was originally applied in medical education and officially introduced as a methodology by Howard Barrows at McMasterUniversity in Ontari o, Canada, in the1960s (Boud &Feletti,1997; Delisle, 1997; Magnussen et al., 2000). The theoretical architecture of Problem-Based Learning (PBL) is rooted in the tradition of experiential learning, which emphasizes that knowledge construction is most effective when situated within meaningful, real-world tasks (Dewey, 1938; Hmelo-Silver, 2004). As Barrows (1996) emphasizes, the value of this approach is not found in memorization, but in providing necessary and effective skills that ensure knowledge remains useful in the future.

The effectiveness of PBL is fundamentally rooted in its organized, iterative tutorial framework, which transitions the emphasis from passive content absorption to active inquiry and the resolution of genuine challenges, allowing students to learn through problem-solving while teachers assume the role of facilitators (Barrows, 2000; Hmelo-Silver, 2004). According to Hmelo-Silver (2004), as shown in Fig.1, this cycle begins with a problem scenario, which prompts students to identify relevant facts and represent the problem. A critical stage in this process is the identification of knowledge deficiencies, termed “learning issues”, which direct the student’s Self-Directed Learning (SDL). By requiring students to formulate hypotheses and subsequently apply newly acquired knowledge to evaluate them, the PBL cycle ensures that learning is grounded in performance.

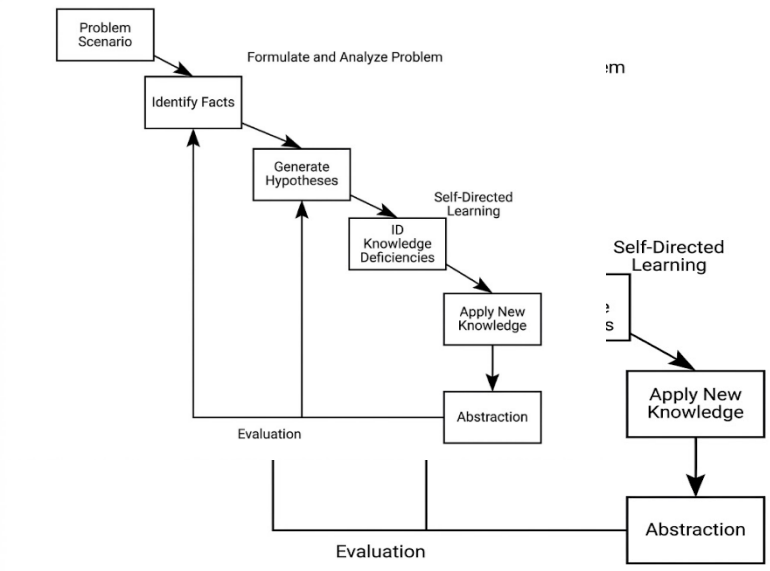


Fig 1. The problem-based learning cycle (Adapted from Hmelo-Silver, 2004)

Vietnam has been using the PBL approach since the 1980s to improve the quality of teaching and learning. Some universities have tried using PBL in certain subjects, and the

results showed that it makes students more active and interested in learning. While PBL is becoming more common in different fields, there is still a need for more research on how to use it to teach specific skills at the university level.

2.2. The Integration of PBL and AI

The combination of PBL and Artificial Intelligence (AI) creates a powerful framework for modern education. According to Novak et al. (2025), while PBL provides a student-centered approach focusing on inquiry and collaboration, AI enhances this process by offering personalized learning experiences, real-time feedback, and data-driven insights. For instance, AI can analyze a student's progress and adjust instructional materials to fit their specific needs, helping them navigate complex problems more efficiently. Similarly, Shocheb et al. (2025) stated that AI can also serve as a virtual mentor that offers guidance, support, and resources while students work through problems. Hence, AI-powered tools, such as intelligent tutors and simulation platforms, can guide learners through every stage of the PBL cycle, from defining the problem to testing hypotheses and reflecting on outcomes (Novak et al., 2025).

Integrating AI into the PBL framework also helps more students and teachers participate in digital learning. According to Omeh and Ayanwale (2025), this combination allows learners to immerse themselves in intelligent environments that offer instant feedback, chances to work together, and automatic help with solving problems. These tools make the PBL learning process more accessible and interactive for everyone involved.

2.3. Impacts on Public Speaking: Motivation, Confidence and Learning Autonomy

AI powered learning enhances accessibility, personalization, and digital inclusion, promoting equitable and quality education through tailored feedback and adaptive learning (Hariyanto et al., 2025). These contributions are particularly relevant in higher education, where AI-supported PBL models can support self-directed learning, critical thinking, and digital literacy.

Public speaking is widely recognized as a fundamental component of professional skill, allowing individuals to persuade others, inspire audiences and boost their confidence. However, for most learners, managing speech anxiety remains the primary obstacle to proficiency. Research indicates that structured instruction and consistent feedback are the most effective methods to reduce nervousness and improve delivery.

As a result, the integration of Artificial Intelligence (AI) has recently opened up new opportunities for speaking practices to address the aforementioned issues in public speaking. Since AI systems enable personalized learning experiences with interaction and feedback, students can engage in repeated practice, correct their own errors, and progressively develop confidence in their speaking abilities (Zhang, 2025). Other research has also proved that AI has had a beneficial impact on the affective outcomes of learners, including the reduction of language anxiety (Bashori et al., 2020; 2021), the enhancement of motivation (Chiu et al., 2023), and the rise in self-confidence. Similarly, a research by Phan (2024) demonstrates that applying AI greatly improves university students' speaking performance while also enhancing their motivation and confidence.

In this study, one of the primary objectives of integrating AI within a PBL framework is to cultivate intrinsic motivation in practicing public speaking, which occurs when learners engage with a task due to personal interest, perceived challenge, or a sense of mastery. Motivation is believed to be significantly amplified when students perceive that the outcome of learning is under their control. Hence, by rewarding profound understanding and

independent action rather than comparative performance, the PBL framework fosters an environment where students are intrinsically motivated (Hmelo-Silver, 2004). Therefore, on an AI-PBL platform, students are provided with a "safety net" where they can manage their own learning direction and practice schedules with real-time feedback.

Furthermore, under the PBL framework, proximate, tangible activities, such as successfully navigating a simulated speech, are more motivating than abstract goals according to Bandura (1997) and can boost the learner's self-efficacy. Junaidi et al. (2020) also claims that AI can improve not only the learners' oral fluency but also their self-confidence.

With regards to self-directed learning, the Problem Based teaching approach was introduced with the desire to integrate language and content, which can facilitate autonomous learning. Moreover, according to Lin and Muenks (2025) and Tozzo et al. (2024), learners' sense of mastery and autonomy may be increased by using AI technologies through adaptive feedback and real-time suggestions. As Hariyanto et al. (2025) noted that students are more likely to remain interested and persistent when they are exposed to information that particularly suits learning speed, personalized learning such as AI-supported systems can therefore promote learner autonomy.

2.4. Research gap

Despite the growing availability of AI tools for public speaking, empirical evidence remains limited regarding their actual impact on a learner's holistic development specifically in terms of fluency, confidence, motivation, and autonomy. In Vietnam, this research gap is even more pronounced; while Problem-Based Learning (PBL) has been studied for decades, implementation has still been largely confined to secondary schools, leaving its potential in higher education under-explored (Phan, 2024). Furthermore, the use of combined AI-supported PBL models for teaching soft skills is still a fairly unexplored area, especially in tertiary contexts. This study addresses these gaps by investigating how an AI-supported PBL framework can support the specific communication needs of university students.

3. Methodology

3.1. Research site

This study was conducted at Ho Chi Minh City University of Technology (HCMUT), Vietnamese National University Ho Chi Minh City (VNU-HCM), one of the leading technical universities in southern Vietnam. The university currently enrolls more than 30,000 students, postgraduates, and doctoral candidates across a wide range of engineering and technology disciplines. This site was selected for its large, diverse student population, which provides a representative sample and facilitates efficient data collection for the study.

3.2. Participants

The participants who took part in the study were students from the Ho Chi Minh City University of Technology (HCMUT). The targets consisted of undergraduate students enrolled in effective communication and English-speaking courses. To minimize potential teacher bias, convenience sampling was used to select two intact classes taught by the same instructor. The experimental group interacted with the AI-supported PBL digital environment, while the control group covered identical course aims using traditional teaching methods.

3.3. Study design

This study employs mixed methods, quasi-experimental research design. Due to rigid academic scheduling at Ho Chi Minh City University of Technology (HCMUT), randomly

assigning individual students to experimental and control conditions was not feasible. Instead, two intact, pre-existing classes were utilized. The independent variable is the instructional method. The experimental group engaged with a Learning Management System (LMS) integrated with Artificial Intelligence (AI) within a Problem-Based Learning (PBL) framework, while the control group received traditional classroom instruction. The dependent variables include objective public speaking performance, measured through fluency and confidence, alongside psychological factors such as learner autonomy, cognitive engagement, and attitudes toward AI technology.

3.4. Data collection procedures

Data were gathered using two primary instruments: objective performance assessments and subjective perception surveys. To measure concrete skill development, students in both groups completed a post-test presentation in the final week of the module, in which the speaking performance results of 48 students in experimental class in semester 252 (2026) were compared with data of the control class about 48 students in the same semester who had not used AI popularly yet. These recorded speeches lasted between three and five minutes. Fluency was objectively measured using automated software to track speaking rate and filler word frequency, while confidence and vocal expression were independently graded by lecturer.

Following the post-test, two sets of questionnaires were also utilized in this study, one for participants who engaged with AI-supported PBL instructions on the LMS platform and one for the other group with traditional teaching methods. Firstly, the questionnaires were sent separately via email to two groups of students with 27 students per group. Students' levels of interest were measured using a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating greater levels of interest, and two open-ended questions in both sets of questionnaires. The survey employed a multiple-choice format, including questions related to students' interest in the two instructional methods. This survey was designed to capture students' subjective experiences regarding their learning autonomy, engagement levels, motivation and overall acceptance of the technology.

3.5. Data processing procedures

Quantitative data

The collected quantitative data were analyzed using statistical procedures appropriate for both paired and independent sample comparisons. After this data-cleaning process, a final sample of 27 valid responses per group was retained for analysis, while the remaining responses were discarded due to corrupt or insufficient data. All performance scores and survey responses were imported into SPSS (Statistical Package for the Social Sciences) software for processing. Descriptive statistics, including means, standard deviations, and standard errors of the mean, were first computed to summarize the central tendencies and distribution of scores within each dataset.

To assess students' levels of interest in the two instructional methods, a Paired Samples T-Test was conducted. The questionnaire comprised 27 items with identical content administered to the same group of 27 students under both conditions, yielding two related sets of scores: d_1 (interest in the traditional public speaking method) and d_2 (interest in the AI-supported PBL method), each ranging from a minimum of 27 to a maximum of 135 points. Because each student contributed one score to each method, the data constituted matched pairs rather than independent samples, making the paired t-test the appropriate procedure for isolating the effect of instructional method on perceived interest while controlling for between-

subject variability. The null hypothesis proposed that no difference existed in students' mean interest scores between the two methods.

To evaluate the effectiveness of AI-assisted practice on students' speaking performance, an Independent Samples T-Test was applied to compare the post-test scores of the Control Group (without substantial AI support) and the Experimental Group (with extensive AI support). As these two groups consisted of distinct, unrelated participants, the data were treated as independent samples. Prior to conducting the t-test, an F-Test on the ratio of two variances was performed to verify the assumption of homogeneity of variance between the two groups, determining whether the subsequent t-test should be conducted under the assumption of equal or unequal variances. Once variance equality was confirmed, the pooled estimator of variance was calculated, and the independent samples t-test was conducted to determine whether the difference in mean performance scores between the two groups was statistically significant or attributable to random variation.

In addition, Cronbach's Alpha will be calculated to verify the internal reliability of the psychological survey scales, ensuring that the items within each section of the questionnaire (e.g., cognitive engagement, learning autonomy, fluency and self-efficacy) consistently measure their intended underlying constructs.

Qualitative data

For the qualitative data derived from the open-ended survey questions, a thematic analysis approach will be employed. Students' written responses detailing the most helpful AI features, their specific learning challenges and their changes after applying either learning methods will be systematically reviewed, coded, and categorized into core themes. This qualitative data serves a vital explanatory purpose, providing deeper context to the quantitative statistics and revealing the mechanisms of how the AI-enhanced learning environment influenced student motivation, self-efficacy and autonomy.

Ethical issues

All participants provide formal institutional consent and are assured that their academic records and personal data will remain strictly anonymous. They will have the right to refuse to participate in the research or withdraw from the study at any stage. The information provided by the students will only be used for research purposes and will not be shared with anyone or any organization other than the researchers. The names of each participant in the report will be a random number from 1 to 48, so their identities will also be kept anonymous. Lastly, all hard copies of data related to our research will be safely stored in a locked file drawer while digital files are password protected.

4. Results

The questionnaire consisted of 27 items with identical content administered to assess students' interest in the two instructional methods. Responses were recorded on a 5-point Likert scale (1 = Strongly disagree, 2 = disagree, 3=neutral, 4 = agree, and 5 = Strongly agree) and converted into numerical scores. Let d_1 and d_2 denote the total converted scores reflecting students' interest in the traditional public speaking method and the AI-supported PBL method, respectively. Given 27 items per method, the minimum and maximum possible values of d_1 and d_2 were 27 and 135, respectively. Boxplots were generated to visualize the distribution of both samples. The boxplot 1 and the results presented in Table 1 indicate that the two datasets exhibit relatively similar distributions. The Posttest scores for the Traditional method (N = 27)

yielded a mean of M = 98.52 (SD = 26.41, SE = 5.08), while the Posttest scores for the AI-supported PBL method (N = 27) yielded a mean of M = 100.63 (SD = 26.96, SE = 5.19).

Boxplot 1

The result of students'interest through online survey

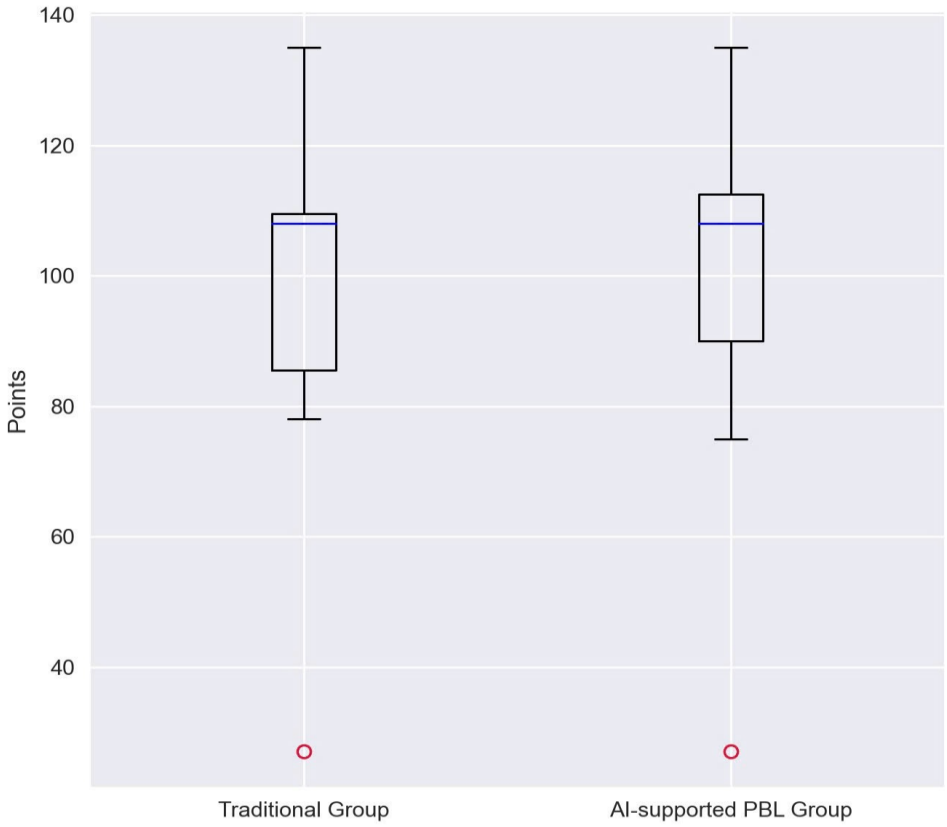


Table 1

Descriptive Statistics for Interest Scores through online Survey

	Group	***N***	***M***	***SD***	***SE***
Online Survey	Traditional	27	98.52	26.41	5.08
	AI-supported PBL	27	100.63	26.96	5.19

Although the descriptive results suggested that the AI-supported PBL sample (d_2) produced a slightly higher mean score than the traditional sample (d_1), this difference could plausibly be attributable to chance. To determine whether the observed difference was statistically meaningful, a paired-samples t-test was conducted, given that both sets of scores were obtained from the same cohort of 27 students, with each participant contributing two related data points.

Let $\Delta d = d_2 - d_1$ represent the difference in converted interest scores between the PBL method and the traditional public speaking method. Three possible outcomes were defined: $\Delta d > 0$ indicates greater interest in the PBL approach; $\Delta d = 0$ indicates equal interest in both methods; and $\Delta d < 0$ indicates greater interest in the traditional approach. The mean difference

score was $\Delta\bar{d}$ = 2.11 (SD = 4.81). The null hypothesis (H_0: $\mu\Delta d = 0$) proposed that students' interest levels did not differ between the two instructional methods. At a significance level of $\alpha = 0.05$ with 26 degrees of freedom, the critical value was t (26) = 1.71. The paired-samples t-test yielded t (26) = 2.24, p-value 0.017, which exceeded the critical threshold and fell below the significance level. The null hypothesis was therefore rejected, indicating that students demonstrated significantly greater interest in the AI-supported PBL method than in the traditional public speaking method, t (26) = 2.24, p-value 0.017.

A second analysis examined the effectiveness of speaking practice with and without substantial AI support, using students' speaking performance scores. The dataset comprised two independent samples: a Control Group, representing speaking practice without significant AI assistance, and an Experimental Group, representing speaking practice with AI support.

As shown in Table 2, descriptive statistics indicated that the Control Group (N = 48) had a mean performance score of M = 6.59 (SD = 0.58, SE = 0.08), while the Experimental Group (N = 48) had a mean performance score of M = 7.25 (SD = 0.76, SE = 0.11).

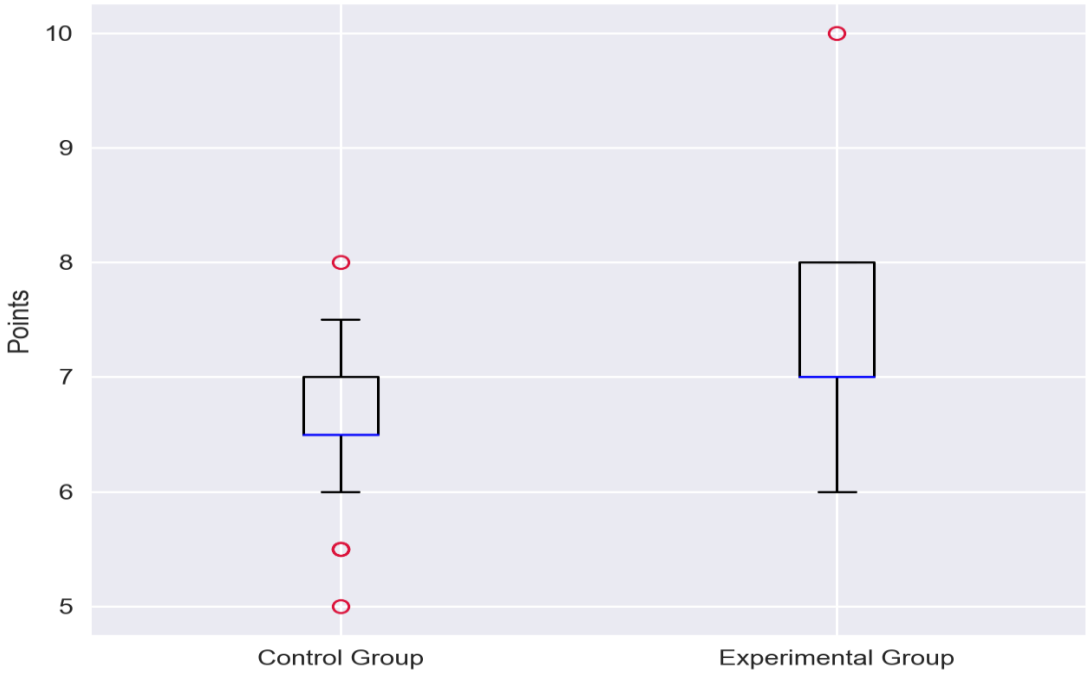
Table 2

Descriptive Statistics for Speaking Performance Scores by Groups

	Group	***N***	***M***	***SD***	***SE***
Posttest	Control	48	6.59	0.58	0.08
	Experimental	48	7.25	0.76	0.11

Boxplot 2:

The speaking performance of groups



Note: The boxplot shows points between control group and experimental group.

The boxplot 2 and descriptive results (table 2) indicated that scores in the Experimental Group were generally higher than those in the Control Group. To determine whether this difference was statistically significant, an independent samples t-test was conducted. Prior to this analysis, an F-test on the ratio of two variances was performed to examine the homogeneity of variance between the two independent populations.

The F-test, conducted at $\alpha = 0.05$, yielded $F = 0.58$, p-value 0.061, indicating no statistically significant difference between the variances of the two populations. Consequently, an independent-samples t-test assuming equal variances was applied. The null hypothesis (H_0) proposed that there was no statistically significant difference between the mean performance scores of the Control and Experimental Groups, and that any observed difference was due to random variation. The pooled estimator of variance was calculated as 0.46.

The independent-samples t-test yielded $t(94) = 4.74$, $p < 0.001$, providing strong evidence to reject the null hypothesis. This result indicated that the Experimental Group's speaking performance scores were significantly higher than those of the Control Group, suggesting that AI support meaningfully enhanced students' speaking skills.

Taken together, the findings from both datasets indicate that students demonstrated greater interest in practicing speaking skills through the AI-supported PBL model relative to the traditional public speaking method, $t(26) = 2.24$, p-value 0.017. Furthermore, students' speaking performance improved significantly when AI-assisted practice was incorporated.

5. Discussion

The findings provide empirical support for the central premise of this study: that embedding AI-driven tools within a Problem-Based Learning (PBL) framework benefits university students' public speaking development across performance, motivation, autonomy, and confidence. Students' stronger speaking performance under AI-supported practice is consistent with Phan (2024) finding that AI integration enhances university students' speaking performance alongside their motivation and confidence, and with Bandura's (1997) self-efficacy theory, which holds that repeated, tangible success in a low-stakes practice environment builds confidence more effectively than abstract instruction alone.

This gain in performance was accompanied by a parallel rise in interest, which can be explained through Hmelo-Silver's (2004) account of PBL as a framework that fosters intrinsic motivation by rewarding independent action and authentic problem-solving rather than comparative performance. Because the questionnaire items were identical across both conditions, this difference can reasonably be attributed to the AI-supported PBL approach itself rather than to extraneous content effects.

This heightened engagement, in turn, points to a plausible gain in learner autonomy: students who found the AI-supported method more interesting were likely also exercising greater control over their own practice and pacing, an arrangement Hmelo-Silver (2004) describes as a "safety net" within the PBL cycle. Taken together, the consistency between greater interest and stronger performance suggests that engagement and skill development are mutually reinforcing within the AI-supported PBL model, lending empirical weight to AI-PBL integration as a pedagogical strategy. While the use of independent samples for the performance comparison limits the strength of causal inference, the convergence of both effects offers reasonable support for the conclusion that AI-supported PBL instruction meaningfully improves students' public speaking outcomes, both in measurable performance and self-reported interest in learning.

Overall, the consistency between the two outcomes—greater interest and superior speaking performance in the AI-supported PBL group—indicates that engagement and skill development are mutually reinforcing within this model. This pattern lends empirical weight to the integration of PBL and AI as a pedagogical strategy and addresses the research gap identified by Phan (2024) regarding the limited empirical evidence on AI integrated PBL implementation in Vietnamese higher education. While the use of independent samples for the performance comparison limits the strength of causal inference, the consistency of both effects provides reasonable support for the conclusion that AI-supported PBL instruction meaningfully improves university students' public speaking outcomes, both in terms of measurable performance and self-reported interest in learning.

Limitations

Despite the promising findings, several limitations should be acknowledged. First, the sample size was relatively small and drawn from a single university context, which may limit the generalizability of the results. Future studies should include a larger and more diverse population across different institutions and cultural settings. Second, the duration of the intervention was relatively short. A longer experimental period may provide deeper insights into the long-term effects of AI-assisted learning on speaking performance and learner autonomy. Third, the study relied partially on self-reported data, which may be subject to response bias. Although efforts were made to ensure anonymity, students may have overestimated their perceived improvements. Finally, while this study focused on speaking skills, other language competencies such as listening and pronunciation accuracy were not examined in depth. Future research could adopt a more comprehensive approach to evaluating language proficiency. These limitations suggest that while the findings are encouraging, they should be interpreted with caution.

6. Conclusion

This study set out to determine the impact of an AI-supported Learning Management System integrated within a Problem-Based Learning framework on university students' public speaking performance, confidence, motivation, and learner autonomy. The findings answer this question clearly: students in the AI-supported PBL condition achieved significantly stronger speaking performance and reported significantly greater interest in their instructional method than their peers in the traditional condition. It also produces a coherent, mutually reinforcing improvement across performance, motivation, and self-directed engagement. This suggests that AI is best treated not as a supplementary tool but as a structural component of the PBL cycle itself.

In addition, the findings highlight the potential of AI technologies to transform traditional language learning environments into more interactive, personalized, and learner-centered experiences. From a pedagogical perspective, the results suggest that educators should consider incorporating AI-driven platforms into their teaching practices to provide students with more opportunities for practice and feedback. Such integration not only improves learning outcomes but also fosters greater learner independence and confidence. Future research is encouraged to explore the long-term impact of AI-assisted learning and to examine its effectiveness across different educational contexts and disciplines.

Acknowledgments

The author would like to express sincere gratitude to all participants for their valuable contributions to this study. Special appreciation is extended to the survey respondents for their time, cooperation, and honest feedback, which were essential to ensuring the reliability of the

research data. The author also thanks colleagues and mentors for their insightful comments and continuous support throughout the research process. The author gratefully acknowledges the organizing committee of the HUFLIT International Conference for providing a valuable academic platform for knowledge exchange and scholarly discussion.

Notes

[1]Nguyen Thi Thao Trang, Foreign Language Center – HCMC University of Technology, *trangnguyen@hcmut.edu.vn*

References

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman and Company.

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: a brief overview. New Directions for Teaching and Learning, 1996(68), 3-12. https://doi.org/10.1002/tl.37219966804

Barrows, H. S., & Barrows, H. (2000). Problem-based learning applied to medical education. Southern illinois university school of medicine.

Bashori, M., Hout, R. V., Strik, H., & Cucchiarini, C. (2020). Web-based language learning and speaking anxiety. Computer Assisted Language Learning, 35(5-6), 1058-1089. https://doi.org/10.1080/09588221.2020.1770293

Bashori, M., Hout, R. V., Strik, H., & Cucchiarini, C. (2021). Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. *System*, *99*, 102496. https://doi.org/10.1016/j.system.2021.102496

Boud, D., & Feletti, G. (1997). The challenge of problem-based learning (2nd ed.). London: Psychology Press.

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with artificial intelligence (AI) based chatbot. *Interactive Learning Environments*, *32*(7), 3240–3256. https://doi.org/10.1080/10494820.2023.2172044

Delisle, R. (1997). How to Use Problem-Based Learning in the Classroom. ERIC Clearinghouse.

Dewey, J. (1938). *Experience and education*. Macmillan.

Hariyanto, Kristianingsih, F. X., & Maharani, R. (2025). Artificial Intelligence in adaptive education: A systematic review of techniques for personalized learning. *Discover Education, 4(1),* Article 32. https://doi.org/10.1007/s44217-025-00908-6

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235–266. https://doi.org/10.1023/b:edpr.0000034022.16470.f3

Junaidi, Hamuddin, B., Julita, K., Rahman, F., & Derin, T. (2020). Artificial intelligence in EFL context: rising students' speaking performance with lyra virtual assistance. International Journal of Advanced Science and Technology.

Lin, S., and Muenks, K. (2025). How students' perceptions of older siblings' mindsets relate to their math motivation, behavior, and emotion: a person-centered approach. Soc. Psychol. Educ. 28:65. doi: 10.1007/s11218-024-09968-2

Magnussen, L., Ishida, D., & Itano, J. (2000). The impact of the use of inquiry-based learning as a teaching methodology on the development of critical thinking. Journal of Nursing Education, 39(8), 360-364.

Novak, L., Kovač, A., & Zupančič, L. (2025). Integrating problem-based learning with artificial intelligence: Transforming student-centered education in the digital era. *International Journal of Educational Technology and Artificial Intelligence, 4(1).* https://topazart.info/e-journals/index.php/ijetai/article/view/84

Omeh, C. B., & Ayanwale, M. A. (2025). Artificial Intelligence Meets PBL: Transforming Computer-Robotics programming motivation and engagement. *Frontiers in Education*, 10. https://doi.org/10.3389/feduc.2025.1674320

Phan, T. N. L. (2024). Investigation into the effectiveness of using problem-based learning in teaching communication skills in English to engineering students. *Interdisciplinary Journal of Problem-Based Learning*, *17*(1). https://doi.org/10.14434/ijpbl.v17i1.33869

Phan, L. T. (2024). Application of AI in enhancing speaking performance of students in Hanoi. *Vietnam Academy of Social Sciences's VSSRA*, *218*(1), 105–128. https://doi.org/10.56794/vssr.1(218), 105-128

Shocheb, A., Ibrohim, I., & Muntholib, M. (2025). The effect of the problem-oriented project-based learning model assisted by artificial intelligence on students' problem-solving and communication skills. *Edubiotik: Jurnal Pendidikan, Biologi dan Terapan*, *10*(02), 424–441. https://doi.org/10.33503/ebio.v10i02.1551

Tozzo, M. C., Reis, F. J. J. D., Ansanello, W., Meulders, A., Vlaeyen, J. W., & Oliveira, A. S. D. (2024). Motivational impact, self-reported fear, avoidance, and harm perception of shoulder movements pictures in people with chronic shoulder pain. European Journal of Physiotherapy, 27(3), 136-146. https://doi.org/10.1080/21679169.2024.2359968

Zhang, X. (2025). Enhancing oral English learning through AI: A case study on the impact of AI-driven speaking applications among Chinese University students. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1595818

Bionote

Ms. Thi Thao Trang Nguyen holds a Master's degree in TESOL through a joint program with Victoria University, Australia. She has over 16 years of experience teaching Non-English-Major learners at Ho Chi Minh City University of Technology and at Language Centers specializing in IELTS and TOEIC. Her research and teaching interests focus on technology-enhanced language learning, soft skills development, and innovation in English education. She has led student teams in award-winning community projects and participated in international professional development programs, including Train the Trainer and online instructional design workshops. She continues to explore innovative approaches to enhance language learning outcomes and foster learner autonomy in higher education contexts.

Appendix A

Experimental Group

Survey on learning experience with ai and the pbl model

Student Information

Student ID: _____

Note: The student ID is used solely to match your performance improvement results. All personal information will be kept strictly confidential.

Instructions: Please indicate your level of agreement on a scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*).

Section 1: Perceived Usefulness of Traditional Learning Methods

1. Instructor feedback in class helps me clearly identify my strengths and weaknesses.

2. Practicing speaking directly with peers/instructors provides a realistic learning environment.

3. Course materials and the instructional structure help me understand the lesson content systematically.

4. In-class practice activities help me learn how to handle real presentation situations.

5. Feedback from instructors and peers helps improve my speaking fluency.

6. I find the current speaking instruction method effective in improving my skills.

Section 2: Cognitive Engagement

7. I feel fully engaged when participating in class discussions or presentations.

8. I spend considerable time reviewing and correcting my mistakes based on instructor feedback.

9. I am willing to rehearse my presentations multiple times to prepare thoroughly for class presentations.

10. In-class activities require me to think carefully about delivering the most persuasive presentation.

11. I remain fully focused and undistracted during speaking practice activities.

12. I proactively seek additional learning materials to improve my speaking skills.

Section 3: Learning Autonomy

13. I set personal goals for each AI practice session (e.g., avoiding reading notes, controlling speaking pace).

14. I monitor my progress by comparing results across different attempts on the system.

15. AI tools enable me to practice effectively without direct teacher guidance.

16. I independently identify ways to correct errors highlighted by AI (e.g., practicing pronunciation until recognized as correct by the system).

17. I feel fully in control of managing my own learning (e.g., choosing when to practice without waiting for instructor instructions).

18. AI evaluation criteria help me assess my speaking performance objectively.

19. I feel comfortable identifying areas for improvement without external reminders.

Section 4: Fluency and Self-Efficacy

20. I feel confident in delivering a fluent presentation in front of a real audience.

21. I believe I can effectively manage presentation anxiety (e.g., nervousness, forgetting content).

22. I am confident in using body language (gestures, movements) to enhance my presentation.

23. I am able to maintain natural eye contact with the audience.

24. I believe I can organize ideas logically and clearly when presenting them.

25. My confidence has increased due to repeated practice in a safe AI-supported environment.

26. I can maintain a stable speaking pace without being too fast or too slow.

27. I am able to minimize filler words (e.g., “um,” “uh,” “you know”) when speaking in public.

Section 5: Technology Acceptance

28. The LMS interface and AI tools are user-friendly and easy to operate.

29. I rarely encounter technical issues that interrupt my practice.

30. I would like to continue using these AI simulation tools in other courses.

Section 6: Open-Ended Questions (Optional – for qualitative analysis)

31. Which feature of this AI system do you like the most? (e.g., virtual audience, detailed feedback reports, ability to repeat practice multiple times, etc.)

32. How has learning through problem-based learning (PBL) combined with AI changed the way you prepare a presentation? (Please briefly describe your changes.)

……

Appendix B

Control Group

Survey on learning experience in public speaking skills

Student Information

Student ID: _____

Note: The student ID is used solely to match your skill improvement results. All information will be kept confidential.

Instructions: Please rate your level of agreement from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*) based on your learning experience during this semester.

Section 1: Perceived Usefulness of Traditional Learning Methods

1. Instructor feedback in class helps me clearly identify my strengths and weaknesses.

2. Practicing speaking directly with peers/instructors provides. a realistic learning environment.

3. Course materials and the instructional structure help me understand the lesson content systematically.

4. In-class practice activities help me learn how to handle real presentation situations.

5. Feedback from instructors and peers helps improve my speaking fluency.

6. I find the current speaking instruction method effective in improving my skills.

Section 2: Cognitive Engagement

7. I feel fully engaged when participating in class discussions or presentations.

8. I spend considerable time reviewing and correcting my mistakes based on instructor feedback.

9. I am willing to rehearse my presentations multiple times to prepare thoroughly for class presentations.

10. In-class activities require me to think carefully to deliver the most persuasive presentation.

11. I remain fully focused and undistracted during speaking practice activities.

12. I proactively seek additional learning materials to improve my speaking skills.

Section 3: Learning Autonomy

13. I set personal goals after receiving feedback from instructors or peers.

14. I monitor my progress by recording myself or practicing speaking at home.

15. I am able to practice effectively without constant supervision from the instructor.

16. I independently identify ways to correct the errors I notice during practice.

17. I feel capable of managing my own learning pace and schedule.

18. I can objectively evaluate my speaking performance based on the criteria provided by the instructor.

19. I feel comfortable identifying areas for improvement without external reminders.

Section 4: Fluency and Self-Efficacy

20. I feel confident that I can deliver a fluent presentation in front of a real audience.

21. I believe I can effectively control presentation anxiety (e.g., nervousness, forgetting content).

22. I am confident in using body language (gestures, movements) to enhance my presentation.

23. I am able to maintain natural eye contact with the audience.

24. I believe I can organize my ideas logically and clearly when presenting them.

25. My confidence has increased as a result of repeated practice during the course.

26. I am confident that I can maintain a stable speaking pace without speaking too fast or too slowly.

27. I believe I can minimize the use of filler words (e.g., “um,” “uh,” “you know”) when speaking in public.

Section 5: Feedback (Optional)

28. Which teaching method or classroom activity has helped you improve your speaking skills the most?

29. What is the biggest difficulty you face when practicing speaking skills using the current learning approach?

#182 STUDENT PERCEPTIONS OF AI-ASSISTED LEARNING AND ITS PERCEIVED ASSOCIATION WITH TEACHER–STUDENT INTERACTION IN EFL SPEAKING CLASSES: AN EXPLORATORY SELF-REPORT STUDY WITH NON-ENGLISH MAJOR UNIVERSITY STUDENTS

Thanh Thuy Nguyen[1]

Abstract

The growing integration of artificial intelligence (AI) in English language education has generated increasing interest in its perceived relationship to classroom interaction. This study presents an exploratory, self-report investigation of student perceptions of AI-assisted learning and its perceived association with teacher–student interaction in English as a Foreign Language (EFL) speaking classes among non-English major university students at HCMC University of Technology. Data were collected from 200 students through Likert-scale questionnaires and reflective journals. The results indicate that students perceived AI tools as associated with increased confidence, greater reported enjoyment of speaking tasks, and more prepared engagement with teachers in specific interactional contexts. Specifically, 64% of students reported feeling more confident after AI-assisted speaking tasks, and 61% expressed that learning through AI felt enjoyable; however, only 47% reported feeling willing to share ideas in class, illustrating a preparedness–participation gap between internal readiness and observable engagement. Notably, challenges such as over-reliance on AI (42% of students) and reduced spontaneity (38%) were also prominent findings and receive equal weight in this analysis. Although most participants expressed positive attitudes toward AI use in speaking classes, its perceived association with interactional quality appears contingent on thoughtful pedagogical integration and appropriate instructional design. All findings are strictly perceptual and associational; no causal claims regarding AI's effects on teacher–student interaction are made. The findings highlight the importance of balancing technological support with active teacher guidance to sustain meaningful classroom interaction.

Keywords: artificial intelligence (AI), AI-assisted learning, teacher–student interaction, EFL speaking, non-English major students

1. Introduction

The rapid advancement of artificial intelligence (AI) technologies has profoundly reshaped English as a Foreign Language (EFL) teaching. From intelligent grammar checkers and pronunciation coaches to conversational chatbots and adaptive learning platforms, AI tools are increasingly embedded in language learning environments worldwide. In Vietnam, where English proficiency is a national educational priority, the integration of AI into higher education classrooms has accelerated considerably, particularly following the COVID-19 pandemic, which forced widespread adoption of digital learning tools (Pham & Ho, 2020).

[1]*Corresponding author. Foreign Language Center, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam. Email: nguyenthanhthuy@hcmut.edu.vn*

Despite this proliferation, a critical question remains underexplored: to what extent do AI-assisted tools genuinely enhance teacher–student interaction in EFL speaking classes, particularly among non-English major students? Speaking is widely regarded as the most challenging language skill, requiring spontaneity, fluency, accuracy, and confidence—qualities difficult to nurture in large, exam-focused Vietnamese university classrooms (Nguyen et al., 2021). Non-English major students, who study English as a compulsory subject rather than a specialized field, often exhibit lower motivation and higher anxiety toward speaking tasks compared to English-major peers (Tran, 2020).

Teacher–student interaction is a cornerstone of effective speaking instruction. Sociocultural theory views second language development as a socially mediated process in which learners' participation in social activity and interaction contributes to their development (Lantolf & Beckett, 2009). When AI tools enter this interaction space, they create new communicative dynamics: students may practice speaking with AI before engaging with teachers; teachers may use AI-generated feedback to redirect classroom discussions; and the feedback loop between teacher, student, and language may be fundamentally altered. Understanding how students perceive these shifts, and what those perceptions suggest for pedagogical practice, is the central motivation of this study.

This paper reports findings from an exploratory study conducted at HCMC University of Technology (HCMUT), involving 200 non-English major undergraduate students enrolled in required EFL speaking courses. Using a mixed-methods design combining Likert-scale questionnaires and reflective journals, the study is guided by three research questions presented in the following subsection.

1.1 Research Questions

The study is organized around three research questions that collectively address AI's perceived influence on teacher–student interaction:

RQ1: How do non-English major university students perceive AI-assisted learning as associated with their readiness and willingness to engage in teacher–student interaction in EFL speaking classes?

RQ2: What interactional challenges do non-English major students report perceiving in AI-assisted EFL speaking activities, and how do they perceive these challenges as affecting their engagement in teacher–student interaction?

RQ3: Under what self-reported pedagogical conditions do students perceive AI-assisted learning as most supportive of their engagement in teacher–student interaction in EFL speaking classes?

1.2 Organization of the Paper

To situate this inquiry within the broader field, Section 2 reviews key theoretical and empirical work on teacher–student interaction in EFL contexts, the role of AI in language learning, and the intersection of these two areas. Section 3 describes the research methodology, followed by the results in Section 4, discussion in Section 5, and conclusions in Section 6.

2. Literature Review

2.1 Teacher-Student Interaction in EFL Contexts

Teacher–student interaction has long been recognized as a central mechanism through which language acquisition occurs. Interaction hypothesis theory (Long, 2015) proposes that

negotiation of meaning between interlocutors drives second language development by making input comprehensible and pushing learners toward more accurate output. In EFL speaking classrooms, this negotiation typically unfolds through question-and-answer exchanges, corrective feedback, recasts, and collaborative tasks (Walsh, 2011).

However, large class sizes, time constraints, and student anxiety frequently impede interaction in Vietnamese university EFL contexts. Research consistently documents teacher-dominated discourse in such settings, where students speak minimally and rely on rehearsed formulaic utterances (Nguyen et al., 2021; Tran, 2020). Interventions that redistribute interactional opportunities, which encourage students to take more turns, elaborate more fully, and respond more spontaneously, are therefore of considerable pedagogical interest.

Against this backdrop of interactional challenges in EFL classrooms, researchers and educators have increasingly turned to AI-powered tools as a potential means of expanding and enriching opportunities for language practice and feedback.

2.2 AI in Language Learning

Artificial intelligence applications in language education encompass a broad spectrum of tools, including natural language processing (NLP)-based feedback systems, speech recognition software, conversational agents, and generative AI models such as ChatGPT (Kohnke et al., 2023; Godwin-Jones, 2022). These tools share a common capacity to provide immediate, individualized feedback—a feature that distinguishes them from traditional classroom-based instruction, where teacher attention is distributed across many learners.

Studies have demonstrated that AI-mediated feedback can reduce speaking anxiety by offering a low-stakes practice environment (Huang et al., 2023). Students reluctant to speak in front of peers may find it easier to experiment with language when interacting with an AI system. Furthermore, AI tools can extend interaction beyond class hours, allowing students to practice pronunciation, vocabulary, and conversational routines independently (Godwin-Jones, 2022). Generative AI tools, particularly large language models, have introduced new possibilities in EFL instruction. ChatGPT and similar tools can generate topic-relevant speaking prompts, simulate conversational partners, provide error analysis, and scaffold content preparation—functionalities that can significantly support speaking task readiness (Kohnke et al., 2023). Nevertheless, concerns about pedagogical dependency, the authenticity of AI-generated content, and equity of access have also been raised (Ulla et al., 2023).

2.3 AI and Teacher–Student Interaction

The intersection of AI and classroom interaction remains a relatively emergent research area. Some studies suggest that AI can serve as an interactional catalyst, meaning students who have prepared with AI tools arrive in class better equipped to participate meaningfully, thereby enriching teacher–student exchanges (Huang et al., 2023). Others caution that AI use may create a paradox: while students interact more with technology, their interaction with teachers may paradoxically diminish as AI becomes the primary feedback provider (Ulla et al., 2023).

This tension underscores the importance of instructional design. When teachers explicitly frame AI tools as preparatory scaffolds rather than replacements for classroom dialogue, positive interactional outcomes are more likely. Conversely, unstructured AI use may result in surface-level language performance that masks deeper communicative deficiencies (Kohnke et al., 2023). Research in Vietnamese EFL contexts is sparse. Pham and Ho (2020) found that students valued digital feedback tools during the pandemic but noted that teacher presence remained essential for motivated participation. No study has systematically explored

student perceptions of AI's relationship to teacher–student speaking interaction among non-English major university students in Vietnam—a gap this exploratory study aims to address.

2.4 Conceptual Framework: The Preparedness–Participation Gap

Recent scholarship on AI in language education presents an emerging paradox. On one hand, AI-assisted tools enhance learner confidence, reduce anxiety, and increase readiness to engage in speaking tasks (Huang et al., 2023; Godwin-Jones, 2022). On the other hand, a growing body of research raises concerns that increased interaction with AI may reduce learners' reliance on teacher-mediated feedback and limit spontaneous classroom communication (Ulla et al., 2023; Kohnke et al., 2023).

To better understand this dynamic, the present study proposes a three-dimensional model of interactional readiness with clear operational distinctions between dimensions: (1) cognitive readiness—operationalized as learners' self-reported ability to generate ideas, organize content, and formulate linguistically appropriate responses prior to speaking tasks, as measured by items related to preparation quality and idea generation; (2) affective readiness—operationalized as learners' self-reported emotional orientation toward speaking, including confidence, perceived self-efficacy, and enjoyment, as measured by items from the anxiety and AI Attitudes subscales; and (3) interactional engagement—operationalized as learners' self-reported willingness to take communicative risks in public classroom settings, including willingness to share ideas, speak without extensive preparation, and engage responsively with teacher prompts, as measured by items from the WTC subscale. Unlike the first two dimensions, the third is inherently social and evaluative, shaped by concerns about peer perception, face-saving, and audience awareness that are not reducible to internal preparedness alone.

This framework suggests that AI may influence these dimensions unevenly: while it is likely to enhance cognitive and affective readiness, its impact on actual interactional engagement may be more limited. This potential discrepancy is conceptualized here as the preparedness–participation gap, defined as the divergence between learners' self-reported internal readiness to communicate and their self-reported willingness to engage in classroom interaction. This concept is related to, but distinct from, the broader willingness to communicate (WTC) construct (MacIntyre et al., 2001) and foreign language anxiety (Horwitz, 2010). WTC captures a learner's general disposition to initiate communication when given the opportunity; research has further linked WTC to self-regulatory and action control processes that shape whether readiness is converted into actual communicative behavior (MacIntyre & Doucette, 2010). The preparedness–participation gap focuses specifically on the mismatch that arises when AI-mediated preparation raises internal readiness without proportionally increasing the willingness to take interactional risks in public, teacher-evaluated settings. Foreign language anxiety, meanwhile, is primarily an affective variable; the gap concept foregrounds the social and contextual barriers that persist even after affective readiness has improved.

3. Research Methodology

3.1 Research Design

A mixed-methods research design was adopted, combining quantitative Likert-scale questionnaires with qualitative reflective journal entries. This triangulated approach enables convergent validation of findings while capturing both the breadth of student attitudes and the depth of individual experiences. The study was conducted during the second semester of the 2023–2024 academic year at HCMUT.

3.2 Participants

Two hundred (200) undergraduate students participated in this study. All were enrolled in required English communication courses representing non-English major disciplines including Engineering, Computer Science, Business Administration, and Natural Sciences. Ages ranged from 18 to 22 years; approximately 53% identified as male and 47% as female. In terms of English proficiency, students were placed into the course based on institutional placement results aligned to the CEFR framework: approximately 41% were assessed at A2 level, 47% at B1 level, and 12% at B2 level. Proficiency level is reported as a descriptive variable; the sample size and study design do not permit meaningful subgroup comparisons across levels, and proficiency is not included as a covariate in inferential analyses. None had previously received formal instruction on using AI tools for EFL learning, though 78% reported prior informal use of translation applications. The eight-week intervention was implemented across four intact class sections, all taught by the same instructor, who is also the author. Holding postgraduate qualifications in Applied Linguistics, the instructor completed several training courses in AI-assisted teaching prior to the intervention, ensuring standardized guidance on AI tool use, task framing, and journal prompt delivery across all four sections.

3.3 AI Tools Used

Students were introduced to two primary AI tools during the eight-week intervention period: (1) ChatGPT (OpenAI) for speaking topic preparation, idea generation, and self-review of drafted oral responses; and (2) ELSA Speak, an AI-powered pronunciation coaching application, for phoneme-level practice and fluency feedback. Teachers provided structured guidance on using these tools as pre-task preparation aids and post-task self-assessment resources, following principles of scaffolded technology integration (Godwin-Jones, 2022).

3.4 Intervention Procedure

The intervention was implemented over eight consecutive weeks during the regular English communication course. Each weekly session lasted approximately 150 minutes and incorporated AI-assisted preparation activities prior to classroom speaking tasks. Students primarily used ChatGPT to generate ideas, organize responses, and rehearse speaking content, while ELSA Speak was used for pronunciation and fluency practice. Teachers emphasized that AI tools should function as preparatory supports rather than substitutes for classroom performance. AI use did not contribute directly to students' course grades.

Table 1

Overview of the Eight-Week AI-Assisted Speaking Intervention

Week	Speaking Focus	AI-supported Preparation	Classroom Speaking Activity	Teacher Role
1	Giving suggestions for improvements	Brainstorm ideas and organize responses (ChatGPT)	Pair discussion and presentation	Introduced AI prompting strategies and provided a speaking outline
2	Talking about the inventions of the 20th century	Search for information, generate key ideas, and organize content (ChatGPT)	Individual speaking and presentation	Guided speaking requirements provided feedback
3	Describing jobs	Generate an outline, key vocabulary, and sample expressions (ChatGPT)	Pair discussion	Facilitated warm-up and follow-up activities

Week	Speaking Focus	AI-supported Preparation	Classroom Speaking Activity	Teacher Role
4	Conducting a job interview	Generate interview questions and model responses (ChatGPT)	Pair role-play	Monitored interactions and provided corrective feedback
5	Discussing health issues	Brainstorm ideas and useful vocabulary (ChatGPT)	Small-group discussion	Facilitated discussion and encouraged spontaneous responses
6	Role-playing a conversation at a doctor's	Generate sample dialogues and useful expressions (ChatGPT)	Pair role-play	Observed role-plays and provided communicative feedback
7	Narrating a folk tale	Brainstorm ideas and organize the story sequence (ChatGPT)	Pair storytelling	Facilitated pre-speaking activities and follow-up discussion
8	Describing favorite entertainment activities	Generate an outline and supporting vocabulary (ChatGPT)	Individual video-recorded presentation	Provided individualized feedback on content and language use

**ELSA Speak was used for pronunciation and fluency practice as homework and is therefore not presented in this table.*

Students spent varying amounts of time using AI tools before each classroom speaking activity. The teacher, who is also the author, followed a standardized instructional protocol developed prior to the intervention in all four classes, including identical guidance on AI use, and speaking task procedures. Students were encouraged to evaluate and personalize AI-generated responses rather than memorizing them. AI-assisted preparation served as a learning support and was not assessed separately from normal course participation.

3.5 Data Collection Instruments

3.5.1 Questionnaire

A 9-item, five-point Likert-scale questionnaire (1 = Strongly Disagree to 5 = Strongly Agree) was administered at the end of the intervention period. Items were selectively adapted from three established, validated scales to ensure psychometric soundness and theoretical coherence across the three dimensions of the study.

The first cluster, measuring speaking anxiety, fear of negative evaluation, and communication apprehension, was adapted from the Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al., 1986). Items were reworded to reflect the AI-assisted speaking context. The second cluster, capturing willingness to communicate, participation intent, and interactional readiness, was adapted from the WTC scale (MacIntyre et al., 2001). The third cluster, addressing student attitudes toward AI tools, perceived enjoyment, and outcome expectations, was adapted from the GenAIAS (Orhan et al., 2024).

The adapted instrument was piloted with 25 non-English major students to verify item clarity and internal consistency. The final questionnaire demonstrated satisfactory reliability (overall Cronbach's $\alpha = .83$; anxiety cluster $\alpha = .79$; WTC cluster $\alpha = .81$; AI attitude cluster $\alpha = .85$), all meeting the conventional threshold of $\alpha \geq .70$ for exploratory research (Tavakol & Dennick, 2011).

In addition, AI-assisted preparation was measured as an independent construct using three items developed specifically for this study (Cronbach's $\alpha = .78$; $M = 3.72$, $SD = 0.61$), serving as the independent variable in the exploratory path analysis reported in Section 4.2.

3.5.2 Reflective Journals

Qualitative data were collected through structured guided reflective journals administered after each AI-assisted speaking task across the eight-week period, grounded in Boud, Keogh, and Walker's (1985) three-stage model of reflection—returning to experience, attending to feelings, and re-evaluating the experience—as more recently synthesized in Mali's (2025) review of reflective practice in language teaching and learning. Following this three-stage model, three structured prompts were developed: (a) "Describe how you used the AI tool to prepare for today's speaking task"; (b) "How did using the AI tool affect how you felt about speaking in class today?"; and (c) "What would you do differently in your next AI-assisted speaking task, and why?" Students were permitted to write in English or Vietnamese to minimize language barriers. Vietnamese-language entries were translated into English by a bilingual research assistant with a postgraduate qualification in Applied Linguistics prior to thematic coding. Back-translation of a randomly selected 15% subsample confirmed semantic equivalence, with no substantive discrepancies identified.

3.6 Data Analysis

Quantitative data were analyzed using descriptive statistics and inferential analyses in SPSS 26.0, including multiple regression to examine whether AI-related affective factors predict willingness to communicate, and exploratory path analysis (with bootstrapping, 5,000 resamples) to explore whether confidence mediates the association between AI-assisted preparation and participation. An exploratory factor analysis (EFA) using principal axis factoring with varimax rotation was conducted to validate the adapted questionnaire's three-factor structure. Given the cross-sectional design, the exploratory path analysis is strictly exploratory; results are presented as associational patterns only and should not be interpreted causally.

Qualitative journal entries were analyzed using Braun and Clarke's (2021) six-phase thematic analysis in NVivo 12. Two independent coders analyzed a randomly selected 20% subsample, yielding inter-rater reliability of $\kappa = .81$, indicating strong agreement (McHugh, 2012). Discrepancies were resolved through discussion.

4. Results

4.1 RQ1: Student Perceptions of AI-Assisted Learning and Perceived Readiness for Teacher–Student Interaction

The findings indicate that students perceived AI-assisted learning as associated with positive shifts in their readiness to engage in classroom interaction, particularly at the level of cognitive and affective preparation. Students reported relatively high levels of enjoyment (61% endorsing Agree or Strongly Agree) and increased confidence after completing AI-assisted speaking tasks (64%). It is important to note, however, that a substantial neutral bloc was present across items: between 31% and 48% of students selected the Neutral response option, indicating that a sizable proportion of participants neither endorsed nor rejected positive associations between AI and their interactional readiness. This neutral bloc is particularly prominent for willingness to share ideas in class (48% Neutral) and expectations about upcoming tasks (40% Neutral), and should temper overly positive readings of the aggregate

endorsement percentages. The Positive (%) column in Table 2 reflects only Agree + Strongly Agree responses; neutral and negative responses together account for the remaining proportion.

Table 2

Affective Readiness and Participation Indicators (N = 200)

Item	SD (%)	D (%)	N (%)	A (%)	SA (%)	Positive (%)
Feel enjoyable	1	3	35	44	17	61%
Have high expectations	1	5	37	42	15	57%
Feel willing to share	2	3	48	34	13	47%
Have no worry	4	10	35	34	17	51%
Feel nervous (anxiety indicator)	2	16	42	25	15	40%†
Feel confident to be seen	6	12	37	27	18	45%
Feel no shame with mistakes	3	7	39	34	17	51%
Feel more confident after tasks	2	3	31	42	22	64%
Expect upcoming tasks positively	3	7	40	33	17	50%

**SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree. Positive (%) = proportion endorsing Agree or Strongly Agree. Percentages are rounded to the nearest whole number and may not sum to 100% due to rounding. † For the negatively worded item "Feel nervous when speaking," higher agreement reflects elevated anxiety rather than a positive outcome. This item is retained in its original directional form to preserve item authenticity; it is not reverse-coded in the descriptive frequency table. In all inferential analyses (regression, mediation, EFA), this item loads on the Anxiety factor as designed and is treated accordingly. The Positive (%) figure of 40% for this item therefore denotes the proportion of students reporting elevated nervousness, not a favorable outcome.*

However, self-reported indicators more directly linked to active classroom participation were comparatively lower. Only 47% of students reported feeling willing to share their ideas in class, and 45% indicated confidence in being seen speaking publicly. These figures should be read alongside the neutral bloc: 48% of students selected Neutral on the willingness-to-share item, and 37% selected Neutral on confidence in being seen speaking, indicating that many students neither endorsed nor rejected participation readiness. Together, these patterns illustrate a perceived gap between positive affective orientation toward AI-assisted preparation and self-reported readiness for live classroom engagement.

The descriptive statistics presented in Table 2 suggest that positive perceptions of AI-assisted learning were stronger for affective readiness than for indicators of public classroom participation. These descriptive patterns were further examined using inferential analyses.

A Pearson correlation analysis revealed a moderate positive association between reported enjoyment and willingness to share in class ($r = .42$, $p < .01$), suggesting that positive affect tends to accompany increased participation. Multiple regression analysis further indicated that confidence is a meaningful predictor of WTC, and exploratory path analysis provides support for confidence as a partial mediator of the association between AI-assisted preparation and WTC. Although causal inferences cannot be drawn from the cross-sectional design, these patterns are broadly consistent with the preparedness–participation gap.

Thematic analysis of the reflective journal entries yielded five overarching themes. Table 3 presents these themes together with frequency counts reflecting the number of individual students (out of 200) whose journals contained each theme across at least one entry.

Table 3

Reflective Journal Thematic Analysis: Themes and Frequencies

Theme	Number of Students (out of 200)	% of Students
Confidence-building	136	68%
Idea generation / content organization	106	53%
Anxiety reduction	94	47%
Improved teacher interaction quality	58	29%
AI dependency concerns	46	23%

[*]Frequency = number of students (out of 200) whose journal entries independently contained each theme across at least one submission round. % of Students = (Frequency / 200) × 100. Note that a single student's journals may contain multiple themes; therefore, percentages do not sum to 100%.

Confidence-building emerged as the dominant theme (68%), reflecting students' repeated accounts of feeling better prepared and less hesitant following AI-assisted preparation. Idea generation was closely related (53%), with many students describing AI as a scaffold for content organization. Anxiety reduction followed (47%), with students noting that AI practice in low-stakes environments helped mitigate fear of public speaking. Improved teacher interaction quality was reported by 29% of students, who observed that their exchanges with teachers became more substantive following AI preparation. Finally, AI dependency concerns emerged as a reflexive theme (23%), with students questioning their reliance on AI-generated content.

4.2 Inferential Analyses of AI-Related Factors and Willingness to Communicate

4.2.1 Multiple Regression Analysis (RQ1)

A multiple regression analysis assessed whether AI-related affective factors, including perceived enjoyment and confidence after AI-assisted preparation, predict students' willingness to communicate (WTC). The overall model was statistically significant, $F(2, 197) = 34.25$, $p < .001$, explaining 25.8% of the variance in WTC ($R^2 = .258$, Adjusted $R^2 = .250$). Confidence emerged as a significant positive predictor ($\beta = .41$, $p < .001$). Enjoyment also reached significance ($\beta = .12$, $p = .047$), though its effect size was considerably smaller than that of confidence.

Table 4

Multiple Regression Predicting Willingness to Communicate

Predictor	B	SE B	β	t	p
(Constant)	1.21	0.28	—	4.32	< .001
Confidence	0.47	0.07	.41	6.71	< .001

Enjoyment	0.14	0.07	.12	2.00	.047

[*]R^2 = .258, Adjusted R^2 = .250, F(2, 197) = 34.25, p < .001. Standardized regression coefficients (β) are reported alongside unstandardized coefficients (B) and their standard errors (SE B).

4.2.2 Exploratory Path Analysis (RQ1)

A exploratory path analysis tested whether confidence mediates the relationship between AI-assisted preparation and WTC. All variables were measured simultaneously in a cross-sectional design; consequently, no temporal precedence can be established among AI-assisted preparation, confidence, and WTC, and the paths described below should be interpreted as associational patterns only. AI-assisted preparation was positively associated with confidence (β = .53, SE = .05, p < .001), and confidence was positively associated with WTC (β = .41, SE = .06, p < .001). The indirect effect of AI-assisted preparation on WTC through confidence was significant based on bootstrapped 95% confidence intervals (β = .22, 95% CI [.14, .31]). The direct effect remained significant after accounting for the mediator (β = .18, SE = .05, p = .001), indicating partial mediation.

Table 5

Exploratory Path Analysis: Indirect Effect of AI-Assisted Preparation on WTC via Confidence

Path	β	SE	95% CI LL	95% CI UL	p
AI-Assisted Prep → Confidence (a)	.53	.05	.43	.63	< .001
Confidence → WTC (b)	.41	.06	.30	.52	< .001
Direct effect (c′)	.18	.05	.08	.28	.001
Indirect effect (a × b)	.22	.04	.14	.31	< .001

[*]Bootstrapped indirect effects are based on 5,000 resamples. CI = confidence interval; LL = lower limit; UL = upper limit. All results are strictly associational given the cross-sectional design.

4.2.3 Exploratory Factor Analysis (Instrument Validation)

An EFA was conducted to examine the factor structure of the adapted nine-item instrument. An important limitation should be acknowledged: with nine items distributed across three factors, the average of three items per factor falls below the four-to-five items per factor recommended in psychometric literature (Costello & Osborne, 2005); EFA results are therefore treated as preliminary. Data suitability indicators were adequate (KMO = .84; Bartlett's Test: $\chi^2(36) = 412.35$, p < .001). Using principal axis factoring with varimax rotation, a three-factor solution emerged corresponding to Anxiety, Willingness to Communicate, and AI Attitudes, explaining 62.7% of the variance. Table 6 presents the factor loadings for all nine items. All items loaded above .55 on their respective factors with minimal cross-loadings (<.30), and factor membership was consistent with the theoretical frameworks from which items were drawn.

Table 6

Exploratory Factor Analysis: Factor Loadings for All Nine Items

Item	Factor 1: Anxiety	Factor 2: WTC	Factor 3: AI Attitudes
Feel nervous when speaking (ANX1)	.74	—	—
Fear making mistakes publicly (ANX2)	.69	—	—
Feel no shame with mistakes (ANX3)	.61	—	—
Willing to share ideas in class (WTC1)	—	.76	—
Speak without extensive preparation (WTC2)	—	.71	—
Feel confident to be seen speaking (WTC3)	—	.65	—
Enjoy using AI tools for EFL (AI1)	—	—	.78
AI improves my confidence (AI2)	—	—	.73
High expectations after AI practice (AI3)	—	—	.67

**Loadings < .40 are suppressed and replaced with "—". KMO = .84. Bartlett's Test of Sphericity: $\chi^2(36) = 412.35$, $p < .001$. Total variance explained = 62.7%. ANX = Anxiety items; WTC = Willingness to Communicate items; AI = AI Attitudes items. All items from the nine-item questionnaire are retained; no items were excluded from this analysis.*

4.3 RQ2: Perceived Interactional Challenges During AI-Assisted Speaking Activities

Despite the positive trends identified above, several challenges emerged that complicate the relationship between AI use and interactional outcomes. Table 7 summarizes the four primary challenges identified through thematic analysis of reflective journal entries, including the proportion of students who reported each challenge and its link to the EFA-validated constructs.

Table 7

Perceived Challenges Affecting Engagement Quality: Student-Reported Frequencies and Implications

Challenge	Students Reporting (%)	Description	Perceived Implication for Engagement
Over-reliance on AI (linked to AI Attitudes factor)	42%	Students submitted AI-drafted responses with minimal personal input	Reduces authentic language production and critical thinking
Reduced spontaneity (linked to WTC factor)	38%	AI preparation led to scripted, less natural oral delivery	Hinders development of real-time communication skills
Digital literacy gaps (linked to AI Attitudes factor)	21%	Some students struggled with AI interface navigation and prompt formulation	Creates unequal learning opportunities in the classroom

Anxiety about AI accuracy (linked to Anxiety factor)	17%	Students uncertain whether AI feedback was pedagogically correct	May undermine trust in self-correction and teacher feedback

[*]Student reporting percentages are derived from thematic analysis of reflective journal entries (N = 200). Percentages reflect the proportion of students who referenced each challenge in at least one journal entry. Factor linkages are based on the EFA structure reported in Table 6.

The most prominent issue was over-reliance on AI-generated content (42%), with students using AI outputs with minimal modification, leading to responses that were linguistically accurate but less personalized and spontaneous. As one student stated, "Sometimes I just copy the answer from AI and try to remember it. When the teacher asks a different question, I don't know how to respond." Closely related was reduced spontaneity (38%), whereby students who relied heavily on AI-prepared scripts expressed discomfort when required to deviate from their responses during live interaction.

Digital literacy gaps affected approximately 21% of participants, who reported difficulty formulating effective AI prompts. As one noted, "I don't know how to ask good questions to ChatGPT, so sometimes the answers are not useful." Finally, 17% of students expressed uncertainty about the pedagogical accuracy of AI feedback, occasionally leading them to disengage from AI-supported preparation.

4.4 RQ3: Self-Reported Pedagogical Conditions Perceived as Supporting Teacher–Student Interaction

Overall, student attitudes toward AI-assisted learning were positive, with a majority indicating willingness to continue using AI tools for English learning. Qualitative findings indicate, however, that positive attitudes were conditional. Table 8 presents the key pedagogical conditions identified through thematic analysis as most likely to support productive teacher–student interaction in AI-assisted contexts.

Table 8

Conditions for Bridging the Preparedness–Participation Gap

Condition	Implementation Strategy	Expected Outcome
Teacher scaffolding	Teachers guide AI use with clear task parameters and explicit framing of AI as preparatory tool	Balanced AI use without over-dependence
Task authenticity	Design open-ended speaking tasks requiring personal opinions and real-time elaboration	Promotes genuine student language production
Digital literacy training	Pre-task orientation on AI tools and critical evaluation of AI output	Reduces inequity in AI access and ability
Reflective practice	Post-task learner reflections on AI use and interaction quality	Deepens metacognitive awareness and language learning

[*]Conditions and strategies are derived from convergent themes across quantitative survey data and qualitative reflective journal entries.

Three key conditions emerged. First, teacher guidance was critical: students who received clear instructions on using AI tools as supports for preparation rather than substitutes for performance reported more productive experiences. Second, task design mattered: tasks requiring personal opinions, real-time responses, or follow-up elaboration were less easily completed through AI alone and therefore encouraged more authentic participation. Third, learner autonomy was important: some students expressed resistance to compulsory or overly structured AI use, suggesting that flexibility may enhance motivation. As one student observed, "AI helps me think faster and gives me ideas I didn't have before, so I feel more confident when preparing."

5. Discussion

5.1 AI as Preparatory Support: Cognitive and Affective Gains

The present study found that non-English major university students perceived AI-assisted learning as primarily associated with the cognitive and affective dimensions of speaking preparation. Students consistently reported that AI tools, particularly ChatGPT for content preparation and ELSA Speak for pronunciation, enabled them to approach speaking tasks with clearer ideas and greater structural coherence, reducing the self-reported cognitive burden associated with spontaneous language production. These perceptual patterns align with Huang et al. (2023), who documented similar anxiety-reduction effects of AI-mediated practice environments in EFL contexts, and with Godwin-Jones (2022), who highlighted the role of AI in extending interactional opportunities beyond classroom hours. From a sociocultural perspective, the reported gains in preparedness may be understood as consistent with the concept of AI as a mediational tool that expands learners' perceived capabilities (Lantolf & Beckett, 2009); however, caution is warranted in invoking the Zone of Proximal Development as an explanatory framework here, as this concept requires the presence of a more capable other and developmental scaffolding across time—conditions not fully verifiable in a cross-sectional, self-report study. The ZPD framing is therefore offered as a tentative theoretical lens rather than a confirmed mechanism.

5.2 The Preparedness–Participation Gap

Despite reported gains in preparedness, self-reported willingness to speak publicly remained comparatively lower—a pattern this study conceptualizes as the preparedness–participation gap. While students reported higher confidence after AI-assisted preparation, indicators more directly linked to active classroom participation, such as willingness to share ideas (47%) and confidence in being seen speaking publicly (45%), were substantially lower. This distinction is consistent with Ulla et al. (2023), who cautioned that AI may be associated with reduced learner engagement with teachers even as technology-mediated practice increases; however, given the self-report nature of the present data, no causal claim can be made about AI reducing interaction. The exploratory path analysis provides associational evidence that confidence partially bridges this gap, but the significant direct effect ($\beta = .18$) confirms that additional factors, likely social, evaluative, and contextual, also shape participation decisions independently of confidence.

Traditional interactional frameworks such as the interaction hypothesis (Long, 2015) emphasize negotiation of meaning as the driver of second language development. The present findings suggest that learners' access to linguistic resources is a necessary but insufficient condition for such negotiation: the willingness to take interpersonal risks in public, evaluative contexts remains a distinct and underexplored barrier. AI tools, which operate in private, low-stakes environments, appear well-suited to the former but less effective in addressing the latter.

5.3 Challenges and the Evolving Role of the Teacher

The study identified four perceived challenges of notable prominence: over-reliance on AI content (42%), reduced spontaneity (38%), digital literacy gaps (21%), and uncertainty about AI feedback accuracy (17%). Critically, the two most frequently reported challenges, namely over-reliance and reduced spontaneity, affect a substantial proportion of participants and warrant equal analytical weight alongside the positive findings. These challenges suggest that AI-supported preparation carries inherent risks for authentic communicative competence when not counterbalanced by appropriate task design and teacher guidance. Specifically, students who submitted AI-drafted responses with minimal personal input reported difficulty adapting when teacher questions deviated from their prepared material, echoing Kohnke et al.'s (2023) observation that unstructured AI use may mask deeper communicative deficiencies. The reduced spontaneity finding (38%) further indicates that scripted preparation, while reducing anxiety, may also constrain the real-time responsiveness that is central to genuine teacher–student interaction. These patterns are associational only; the cross-sectional, single-group design precludes causal attribution of reduced spontaneity to AI use per se.

Rather than diminishing the importance of teachers, these findings suggest that AI integration reconstitutes instructional responsibilities toward the management of interactional and affective conditions. As AI tools assume a greater role in supporting content preparation, teachers become increasingly central in facilitating real-time meaning negotiation, modeling spontaneous language use, and cultivating the social conditions necessary for risk-taking in public communication. This reframing of teacher roles has direct implications for professional development: EFL instructors in AI-integrated contexts may benefit from targeted training not only in AI tool operation, but in designing post-AI interaction tasks, giving metacognitive feedback on scripted versus authentic communication, and managing the socio-affective dynamics that emerge when students arrive over-prepared but under-confident in live interaction.

5.4 Implications for Pedagogy

The central pedagogical finding of this study is that students perceived AI-assisted preparation as most beneficial when embedded within a structured instructional framework rather than used independently. Students who received clear guidance on using AI tools as preparatory scaffolds reported more substantive and confident engagement with teachers, while those who relied on AI outputs without critical evaluation tended toward scripted delivery and reduced responsiveness in real-time interaction. This pattern is consistent with Kohnke et al. (2023), who argued that pedagogical framing is the key variable determining whether AI functions as a communicative bridge or a communicative barrier, and with Ulla et al. (2023), who identified teacher role clarity as essential in AI-integrated classrooms. The explanation for this pattern likely lies in the distinction between AI as a content provider versus AI as a thinking scaffold: students who were guided to use AI for idea generation and then elaborate independently arrived in class better equipped to deviate from their prepared material. This distinction has direct implications for instructional design. Specifically, the present findings support the following evidence-informed strategies: designing speaking tasks that require personal opinions and real-time elaboration, rendering AI-generated scripts insufficient; providing structured low-stakes practice sessions to build social and evaluative confidence alongside linguistic readiness; offering explicit digital literacy training so that all students can access AI tools effectively and equitably, addressing the participation gap identified among the 21% who reported difficulty with AI interfaces; and framing AI use through reflective journal practice that builds metacognitive awareness of when AI assistance enhances versus constrains

authentic communication. These implications extend beyond classroom practice to curriculum design and teacher professional development, where preparation for AI-integrated speaking instruction should be considered a core competency.

6. Conclusion

6.1 Summary of Findings

The findings reveal that students perceived AI-assisted learning as associated with nuanced and uneven shifts in their readiness for teacher–student interaction in EFL speaking classrooms. All findings are self-reported and strictly perceptual; no causal claims about AI's effects on interaction are made. Students broadly perceived AI as associated with increased enjoyment (61%), improved preparedness, and higher post-task confidence (64%). At the same time, a sizable neutral bloc (31%–48% of students) selected the Neutral response across items, indicating that a substantial proportion of participants neither endorsed nor rejected positive associations. The central contribution of this study lies in identifying a preparedness–participation gap: students reported that AI tools were associated with their ability to generate ideas and rehearse language prior to speaking, yet this perceived preparedness did not translate consistently into self-reported willingness to engage publicly in live classroom interaction. Equally important, the study identified significant perceived challenges: 42% of students reported over-reliance on AI-generated content, and 38% reported reduced spontaneity in oral delivery—findings that qualify the overall positive picture and warrant serious consideration in pedagogical planning. Students continued to report hesitation and anxiety when engaging in live classroom interaction, suggesting that AI is perceived as exerting a more limited influence on the socio-affective dimension of public communication.

Effective teacher–student interaction in EFL contexts cannot be fully supported through technological means alone, as interaction depends not only on linguistic readiness but also on confidence, identity, and social risk-taking. As AI increasingly supports the cognitive aspects of language learning preparation, the teacher's role appears even more central in shaping the affective and interactional environment of the classroom. Effective AI integration therefore calls for a balanced instructional approach pairing AI-supported preparation with pedagogical strategies that promote real-time communication, reduce fear of public speaking, and foster authentic participation—while remaining attentive to the dependency and spontaneity challenges that AI integration may introduce.

6.2 Contributions of the Study

This research makes three contributions. First, it introduces the concept of the preparedness–participation gap as a framework for understanding the perceived differential impact of AI on internal readiness versus reported engagement in interaction. Second, it provides exploratory mixed-methods evidence from a Vietnamese university context, a setting underrepresented in research on AI in language education, documenting how non-English major students perceive AI-assisted speaking preparation. Third, it highlights the continued centrality of teacher mediation in AI-integrated classrooms, with implications for pedagogy, curriculum design, and teacher professional development.

6.3 Limitations and Future Directions

Several limitations frame the scope of interpretable findings. First, and most fundamentally, the study relied exclusively on self-reported perceptual data. All conclusions pertain to what students believed about their experiences, not to observable changes in speaking behavior or interaction quality. Future research should incorporate behavioral measures such

as discourse analysis and turn-taking counts to validate perceptual reports. Second, the absence of a control group or pre-test/post-test design means that no baseline comparison is available; observed patterns may reflect pre-existing dispositions rather than responses to the AI intervention. Third, the single institutional setting at HCMUT and the eight-week timeframe limit generalizability to other Vietnamese university contexts, other EFL systems, or higher-proficiency learners. Fourth, the nine-item questionnaire yields an average of three items per factor, falling below the four-to-five items per factor recommended by Costello and Osborne (2005); the EFA results should therefore be treated as preliminary and require replication with a larger, psychometrically developed instrument. Fifth, the cross-sectional exploratory path analysis cannot establish temporal precedence among AI-assisted preparation, confidence, and WTC; the "mediation" model is associational only. Sixth, the study did not examine proficiency as a moderating variable; it is plausible that the preparedness–participation gap manifests differently across A2, B1, and B2 learners, a question that warrants targeted future investigation. Seventh, social desirability bias may have inflated positive self-reports, given that AI integration was teacher-initiated and journal entries were submitted to the instructor. These limitations collectively frame this study as exploratory and call for methodological triangulation in future work.

Future research should incorporate direct behavioral measures of classroom interaction, including discourse analysis and turn-taking frequency counts; longitudinal designs tracking communicative development over time; experimental or quasi-experimental designs for stronger causal evidence; and investigation of teacher perspectives on AI integration. Additional research is also needed to examine how the preparedness–participation gap varies across proficiency levels, task types, and cultural contexts.

References

Boud, D., Keogh, R., & Walker, D. (Eds.). (1985). *Reflection: Turning experience into learning*. Kogan Page.

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1–9. https://doi.org/10.7275/jyj1-4868

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, *26*(2), 5–24. https://doi.org/10.64152/10125/73474

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*(2), 154–167. https://doi.org/10.1017/S026144480999036X

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112–131.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lantolf, J. P., & Beckett, T. G. (2009). Sociocultural theory and second language acquisition. *Language Teaching, 42*(4), 459–475. https://doi.org/10.1017/S0261444809990048

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.

MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition*, *23*(3), 369–388. https://doi.org/10.1017/S0272263101003035

MacIntyre, P. D., & Doucette, J. (2010). Willingness to communicate and action control. *System, 38*(2), 161–171. https://doi.org/10.1016/j.system.2009.12.013

Mali, Y. C. G. (2025). Reflective practice in language teaching and learning: a literature review. *Journal of English Teaching and Learning Issues*. https://doi.org/10.21043/jetli.v8i2.25463

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

Nguyen, L. T., Hung, B. P., Duong, U. T. T., & Le, T. T. (2021). Teachers' and learners' beliefs about pronunciation instruction in tertiary English as a foreign language education. *Frontiers in Psychology*, *12*, Article 739842. https://doi.org/10.3389/fpsyg.2021.739842

Orhan, A., Yıldız, T. A., & Yağcı, Ş. Ç. (2024). Assessing EFL learners' attitudes on generative artificial intelligence: development and validation of generative artificial intelligence

attitude scale for EFL learners (genaias). *Journal of Research on Technology in Education, 58*(3), 494-514. https://doi.org/10.1080/15391523.2024.2437744

Pham, H. H., & Ho, T. T. H. (2020). Toward a 'new normal' with e-learning in Vietnamese higher education during the post COVID-19 pandemic. *Higher Education Research & Development*, *39*(7), 1327–1331. https://doi.org/10.1080/07294360.2020.1823945

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tran, T. Q. (2020). EFL students' attitudes towards learner autonomy in English vocabulary learning. *English Language Teaching Educational Journal*, *3*(2), 86–94. https://doi.org/10.12928/eltej.v3i2.2374

Ulla, M. B., Perales, W. F., & Busbus, S. O. (2023). "To generate or stop generating response": Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand. *Learning: Research and Practice, 9*(2), 168–182. https://doi.org/10.1080/23735082.2023.2257252

Walsh, S. (2011). *Exploring classroom discourse: Language in action*. Routledge.

Author Bionote

Thanh Thuy Nguyen (nguyenthanhthuy@hcmut.edu.vn) is a lecturer at the Foreign Language Center, Ho Chi Minh City University of Technology (HCMUT), Vietnam. She has accumulated over 15 years of experience in English language teaching since joining HCMUT in 2008. She holds a Bachelor's degree in English Language Teaching from Ho Chi Minh City University of Pedagogy and a Master's degree in Applied Linguistics from Curtin University, Australia. Her research interests include AI-assisted language learning, speaking instruction, and teacher–student interaction in EFL contexts. Guided by her philosophy that "teaching is a dual learning process," she believes teachers and students grow together through continuous learning and reflection.

#184 ENHANCING STUDENT ENGAGEMENT IN INTERPRETATION COURSES THROUGH AI-ASSISTED SELF-PRACTICE

Dang Thi Anh Tuyet[1]

Abstract

In interpretation courses, students are typically expected to engage in independent practice on their own to develop these complex skills. However, many students encounter difficulties in practicing interpreting autonomously. The study examined how Artificial Intelligence (AI) integration can facilitate students' self-practice in interpretation courses. This research adopted a qualitative-dominant mixed-methods design, incorporating supportive quantitative analysis to gain better insights into the research questions. The study was conducted with the participation of eighty-two third-year students from three business interpretation classes at the Faculty of Foreign Languages, HUFLIT. The participants were encouraged to utilize AI tools to support their independent practice for self-study. The findings reveal that the majority of students perceived AI-assisted self-practice as both useful and engaging while also offering several recommendations for improving its implementation. For more generalizations and pedagogical implications, it is recommended that further studies on a large-scale and on the negative aspects of the approach should be conducted to examine the effectiveness of AI as a facilitator in self-regulated practice for interpreting skills among university students.

Keywords: engagement, AI integration, interpretation courses, interpreting training, self-practice

1. Introduction

1.1. Background of The Study

With the application of Artificial Intelligence (AI) in all aspects of life, necessary transformations have taken shape in education to adapt to the new trends. AI integration in higher education is currently not only a trend but also a requirement for educators with a view to optimizing learning outcomes. Tertiary education has recently experienced vivid AI integration in many fields since AI applications have proved to benefit teaching and learning considerably, saving both time and effort. In terms of motivation, AI-powered platforms have supposedly enhanced students' engagement and autonomy that are crucial for improvement. Furthermore, with the popularity of AI applications in recent years, students' desire for the freedom of utilizing AI in academic work has risen, which has obliged educators to take it seriously as to how to integrate AI into pedagogical approach effectively.

Business Interpretation course is among the author's favorite subjects at HUFLIT Faculty of Foreign Languages (FFL). However, this module can sometimes be seen as a "nightmare" for many students of different majors, for this subject requires a comprehensive competence in many factors with oral communication skills and cognitive processing reflexes as the prerequisites. Students are supposed to carry out regular practice to hone their interpreting skills on their own. Therefore, students with low levels of proficiency usually find this subject immensely challenging. Before AI applications came into life, students mostly had to practice alone or occasionally with peers. This is indeed demanding work which requires tremendous effort and motivation; otherwise, giving up half-way through may result in failure,

[1] *Lecturer, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Ho Chi Minh City, Vietnam*

giving rise to students' intimidation in learning this subject. Under such circumstance, students supposedly need some new kind of motivation for self-study. In seeking new inspiring and engaging ways to facilitate students' self-practice, the author had a burning desire to take advantage of AI integration. Nevertheless, interpretation courses have intrinsic features that require particular integration approach as the ultimate learning outcomes of this subject is learners' capability to perform their genuine interpreting skills.

1.2. Research Problem

In interpreting training, reality has showed that "practice makes perfect" significantly necessitates students' improvement; what matters is how to engage students to practice interpreting skills frequently. In order to enhance engagement, students would be encouraged to practice with peers in addition to practicing independently. However, practicing with peers is sometimes hindered by several reasons including the matters of time-matching, high pressure from peer feedback. At the outset, the use of AI in interpreting training was mostly restricted, but it has gradually proved to be beneficial in many aspects and has become increasingly applicable. For this study, in empowering students with AI integration into their self-practice in interpretation courses, the author wondered how the approach can benefit students and enhance their engagement.

1.3. Research Objectives

In allowing AI integration into self-practice in interpretation courses, it is hoped that students' self-regulated learning is recognized, thereby helping promote their engagement and motivation. In alignment with this view, this study aims to investigate the efficacy and pedagogical impacts of AI integration on students' engagement in learning interpretation whereby pedagogical implications and recommendations can be proposed for further implementation.

1.4 Research Questions

With the research objectives in mind, my research questions were designed accordingly as follows:

1. *To what extent do students find AI assistance engaging and useful for their self-practice in interpretation courses?*
2. *What are students' perceived benefits from using AI tools to support their self-practice in interpretation courses?*
3. *What are common challenges students may encounter while using AI support for self-practice in interpreting training, and their recommendations for improvement?*

1.5 Significance of The Study

The wave of AI approaching and effecting tertiary education was welcomed as a supplementary drive for boosting students' engagement. With more and more students relying on AI for their translation and even interpretation regardless of not being permitted, I would rather openly allow them with guidelines for proper use than leave them with hidden misleading manner. Therefore, empirical findings from this study can provide insights into the efficiency of AI integration in enhancing learner engagement in interpreting training, thereby hopefully contributing to subsequent implications and pedagogical recommendations for further studies.

2. Literature Review

With the above-said core research questions, the investigation from former researcher on this practice was carried out. In light of the literature, the rapid wave of Artificial Intelligence (AI) has significantly impacted language teaching and learning, prompting novel approach for interpreter training. Interpreting competence requires more than just one element; it in fact involves an intensive cognitive process and particular skills like note-taking, memorizing, etc., of which continuous rigorous practice really matters. In this aspect, AI-assisted self-practice has emerged as a highly welcoming and promising addition for students' self-training. This literature review has drawn on recent papers on the enhancement of interpreting skills through AI supports. It examines the topic through four key themes: current trends, practical applications, pedagogical benefits, and possible challenges.

2.1 Trends in AI Integration in Interpreting Training

In recent years, Translation and Interpreting (T&I) education has undergone paradigm transformation. At the outset, AI was viewed as a disruptive threat that was not welcomed and encouraged for use in T&I pedagogy. However, AI has now been embraced as a pedagogical necessity. Cui et al. (2025), and Šveda & Poláček (2025) highlighted that modern interpreter training must take the initiatives to adapt to the "era of AI," moving from traditional methodologies to integrating technological competencies directly into the curriculum. This trend is echoed globally; for example, Tran (2026) noted the popularity of using AI translation and interpreting tools among Vietnamese students, indicating a dramatic shift where students independently and eagerly employ AI tools for linguistic support and practice. Besides that, as Gutiérrez (2025) put it, the gap between academic training and professional reality has now become mitigated. AI assistance in interpreting teaching and learning is increasingly welcomed as valuable, supportive tools for future careers instead of a threatening mechanism for job displacement. In line with the new trends, educators are prompted to align pedagogical practice with these new technological applications.

2.2 The Application of AI Tools for Enhancing Interpreting Practice

The application of AI tools in training interpreting skills has proved to be popular in various ways. Particularly, with generative AI chatbots, AI can be utilized as simulated interlocutors and automated feedback providers. For self-guided practice, AI can perform various interactive and analytical modalities to simulate real-world interpreting environments. With regard to interactivity, Yu et al. (2025) examined how generative AI chatbots have been employed to deliver targeted critique in comparison with using traditional peer-review mechanisms. Likewise, Chien, Chan, and Hou (2024) noted the efficiency of using generative AI as scaffolding for role-playing educational games. With AI simulating consecutive or simultaneous interpreting environments outside the classroom, giving instant feedback on students' performance, learning outcomes can significantly improve. Going beyond conversational interaction, AI tools can also performs tasks analytically. According to Tang, J. (2023), with AI-based needs analysis in digital environments, self-practice materials can be tailored for specific linguistic deficits or professional domains whereby student's independent study regimen in interpretation can be personalized.

2.3 The Benefits of AI-Assisted Learning in Interpreting

According to researchers, the pedagogical benefits of AI assistance in self practice primarily involve enhancing self-regulated learning, providing instant feedback, and promoting student motivation. High levels of student satisfaction is reported by Consuelo Sáiz-Manzanares et al. (2023) when self-regulated learning is facilitated by chatbots. Similarly, Ng,

Tan, and Leung (2024) claimed that students' autonomy could be empowered with popular platforms like ChatGPT during their educational journeys. Particularly in interpreter training, the availability of 24/7 instant feedback from AI tremendously benefit students in terms of refining their delivery and accuracy while independent and constant practice is essential for cognitive process reflex (Chien et al., 2024; Yu et al., 2025). Interestingly, Yu, Wei & Chen, 2025 pointed out that AI feedback strategies can also help interpreting students identify semantic errors, adjust their grammar structures, and eventually refine their overall performance at their own pace. Notably, in terms of motivation, Tran (2026) pointed out that the these tools can be integrated to positively affect learners' motivation. Specifically, students' anxiety can be reduced, and confidence be increased thanks to immediate scaffolding provided by AI in a low-pressure environment.

2.4 Challenges and Limitations

In spite of profound benefits of integrating AI into interpreting training, notable challenges and limitations also come alongside. One of the most striking limitations is the shortcoming of technology in replicating the nuanced human cognition which can be required in complex, dynamic communicative contexts. As argued by Gómez & Pinazo (2026), in in high-stakes contexts that require cultural mediation, empathy, and situational awareness, AI show incapability in replacing humans due to its inherent deficiencies. Tang, L. (2025) also pointed out "other minds" - an intuitive, empathetic understanding of human intention that AI lacks whereas this cognitive capability is vital for coherence interpreting. The most note-worthy challenge in terms of pedagogy are concerns about learner real autonomy which is a common concern among educators in integrating AI in other domains as well. According to Darvishi et al. (2024), student agency can be negatively impacted if over-reliance on AI assistance is committed. As frequently warned by other researchers, over-reliance on AI support can lead to passive learning and degraded critical thinking skills instead of active cognitive engagement.

Eventually, researchers emphasized that integrating "critical AI literacy" into interpreter training is a urgent necessity. The ultimate goal is to ensure that while using AI for self-practice to possess the skills, students can also critically assess AI outputs, avoid hallucinations, and fully be aware the ethical implications of the tools they use (Giustini & Dastyar, 2024). Despite multiple benefits of utilizing AI assistance in self-practice for interpreting training such as accessible simulations, personalized feedback, and enhanced learner motivation, learners should not consider AI applications a 'magic cure-all'. There should be a balanced pedagogical approach. Leveraging the practical applications of AI as well as actively fostering critical AI literacy should be catered in order to address ethical challenges thereby preserving student agency, and navigating the inherent drawback in machine cognition.

3. Methodology

3.1. Research Design

With an interpretative epistemological stance, this research primarily adopted qualitative methodology; in addition, to evaluate the extent of student engagement and perceived utility, quantitative measures were also incorporated. Overall, this research employed a qualitative-dominant mixed-methods design, incorporating supportive quantitative analysis to gain better insights into the research questions.

3.2. Study Setting

This study was conducted at the Faculty of Foreign Languages (FFL) at HUFLIT with regard to Business Interpretation course which serves as a significant module for students of different majors. This course is considered advanced, requiring comprehensive proficiency in various aspects, such as good speaking and listening skills, vocabulary and general knowledge on daily life and business settings. In addition to independent practice and practice with peers in a traditional manner, the students were allowed to use AI-powered platforms to enhance their engagement for self-practice. In this study, my pedagogical intent is utilizing AI as supportive tools for students' self-practice. Students were intentionally empowered with AI support for their self-practice so that engagement can be promoted. Accordingly, students were encouraged to consult with AI-powered platforms during self-practice. The extent of AI utilization is supposedly limited around mediating functions such as generating learning materials (e.g. a text or speech), transforming text-to-speech, composing a vocabulary list, doing role-playing, giving feedback for interpretation work. The restriction on AI facilitation is definitely set: interpreting skills must be performed by students themselves.

3.3. Participants

The participants in this study were 82 third-year students from three Business Interpretation classes at HUFLIT FFL, being randomly selected by volunteering with informed consent in alignment with institutional ethical guidelines.

3.4. Research Instruments

In order to ensure robust data collection, gaining insights into instructional efficacy with regard to enhancing student engagement in interpretation courses through AI integration, the following three research instruments were employed:

3.4.1. Structured Classroom Observations

Observations were conducted across the nine class sessions using a checklist to document student engagement and behavioral responses during AI-assisted self-practice.

3.4.2. Student Survey:

Informed by preliminary observational data, a structured questionnaire comprising 5-point Likert scale questions, multiple-choice questions, and open-ended question(s) was administered to explore student perceptions and experiential outcomes. The online survey questionnaire was sent to students via Google Forms.

3.4.3. Semi-structured Interviews

Semi-structured focus group interviews were conducted with a sample of nine student volunteers; the interviews lasted approximately five minutes each to examine learner attitudes, identify perceived benefits and constraints, and to gather pedagogical recommendations regarding AI mediation in interpreting self-practice.

Overall, these data collection instruments provided a comprehensive, multi-perspective data, yielding deep insights into the dynamics of the contemporary pedagogical practices.

3.5. Procedure

The study was conducted over a nine-week period during the academic term, including three phases:

3.5.1. Preparation Phase

Guidelines were provided to students, showing how to use AI support for self-practice. The ultimate goal of this approach was clarified at the outset - using AI assistance to improve students' independent interpreting performance. Students should be fully aware that AI-powered tools can only serve as facilitators for their self-practice and that AI can by no means be utilized as an interpreter in lieu of students themselves. To ensure compliance, the instructor stressed the importance of maintaining cognitive autonomy, warning of the adverse effects of AI over-reliance on critical thinking and analytical reasoning and that students' performance from this practice can be appreciated as their efforts, but not be evaluated as formal testing results in class. The awareness of this rule necessarily prevents the students from intentional cheating or over-reliance on AI in their self-practice.

3.5.2. Implementation Phase

Students were encouraged to experience the approach - AI integration in their self-practice, both in class and in self-study, following the provided guidelines. Recommended AI supportive activities basically involve creating materials (e.g. a text or speech), transforming text-to-speech, composing vocabulary list, doing role-playing, giving feedback for interpretation job.

3.5.3. Data Collection Phase

Class observations were carried out during the 9 weeks with field notes being documented while the students engaged in self-practice with AI facilitation. Prominent trends identified during the observations guided further empirical inquiry. Specifically, salient themes derived from these observations facilitated the development of a structured questionnaire including two 5-point Likert scale questions, two multiple-choice questions, and one open-ended question to elicit participant perceptions and qualitative feedback. The survey was administered electronically via Google Forms nearing the end of the course. For the last two week, semi-structured interviews incorporating open-ended questions were conducted in person. All interview sessions were recorded with prior participant consent and subsequently transcribed verbatim for analytical processing.

3.6. Data Analysis

Data analysis was conducted based on qualitative, thematic analysis approach. Accordingly, qualitative data from open-ended responses from online surveys and interviews were analyzed. Transcripts from the interviews were coded to identify recurring themes. Besides, the quantitative data of the first two Likert scale questions in the survey questionnaire were also analyzed using SPSS software platform to strengthen the reliability of the data. Meanwhile, notes from class observations were used to triangulate findings, increasing the validity of interpretations.

In particular, thematic analysis of the interview transcripts revealed several emergent categories: (a) the engagement and utility of AI applications as learning scaffolds, (b) perceived benefits, including personalized learning experience and exposure to more topics for practice, (c) common challenges including student over-reliance and misuse in employing the approach, and (d) student recommendations for optimizing AI integration in interpretation courses. Collectively, these themes offered critical insights into student dispositions, highlighting both the benefits and constraints of embedding AI technologies in interpreting training. Furthermore, quantitative and qualitative data obtained through Google Forms was analyzed to discern broader trends regarding student attitudes and perceptions. The integration of close-ended Likert scales alongside open-ended qualitative prompts provided both statistical breadth and contextual depth to the findings. Ultimately, this mixed-methods analytical framework

yielded triangulated insights into the efficacy of AI support in enhancing learner engagement for self-practice in interpretation courses.

4. Results

4.1 Addressing Research Question 1:

Students' Showing Interest and Confirming the Usefulness of Utilizing AI to Assist Their Self-practice in Interpretation Courses

Table 1

Item Statistics

	Mean	Std. Deviation	N
Q1	4.02	.785	82
Q2	4.20	.744	82

**Note: Q1: Question 1, Q2: Question 2 in the survey questionnaire*

Figure 1

Students' expressing engagement in using AI support for their self-practice in interpretation courses

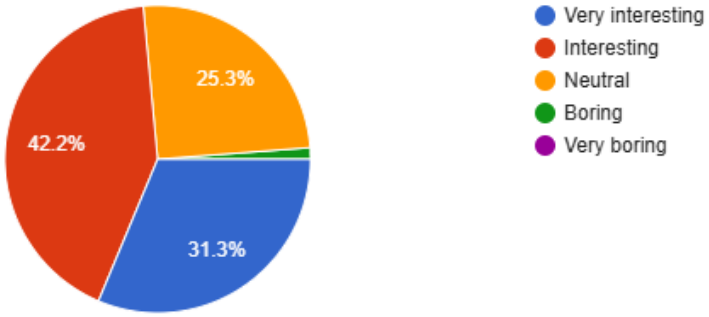


Figure 1 shows that a high proportion of participants welcome the approach, taking interest in utilizing AI to support their self-practice in interpreting training. The majority of students claimed their engagement with AI integration into their self-study as a facilitator; specifically, approximately 75% of students found the approach *interesting* and *very interesting*. This phenomenon accords with the data generated from classroom observations where most of them showed high level of interest and engagement during their self-practice, interacting with AI tools as mediators.

The data reliability can be reinforced in quantitative analysis (in Table 1): for Q1, a standard deviation of 0.785 is relatively low for a 5-point scale, showing a strong consensus among the cohort. Most responses are categorized around the *"Interesting"* and *"Very interesting"*. A mean score of 4.02 indicates that the average respondent finds using AI tools to facilitate interpreting skills to be *"Interesting"*.

Figure 2

High level of usefulness of utilizing AI-powered platforms in practicing interpreting skills

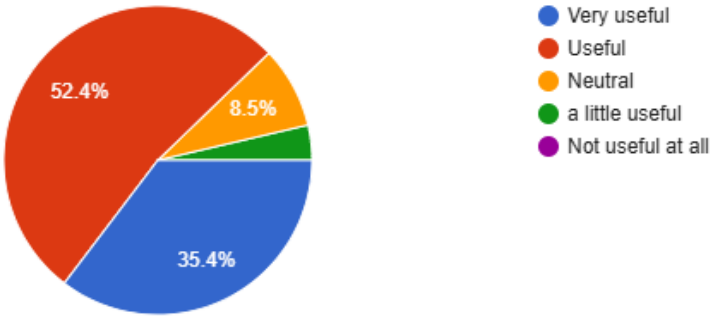


Figure 2 presents students' perceived usefulness of AI tools in interpreting training with regard to usefulness of employing AI as mediator and facilitator in practicing interpreting skills. Impressively, 100% of the participants confirmed the approach to be *useful*; the majority claimed the high level of usefulness of this approach with almost 90% rating as *useful* and *very useful* while no one found it not useful at all. These proportions align with the research findings from the literature review that AI can benefit interpreting training in many ways including enhancing self-regulated learning, providing instant feedback, and promoting student motivation. Additionally, AI mediation can also help identify semantic errors, adjust grammar structures, and eventually refine overall performance at one's own pace, enhancing autonomy.

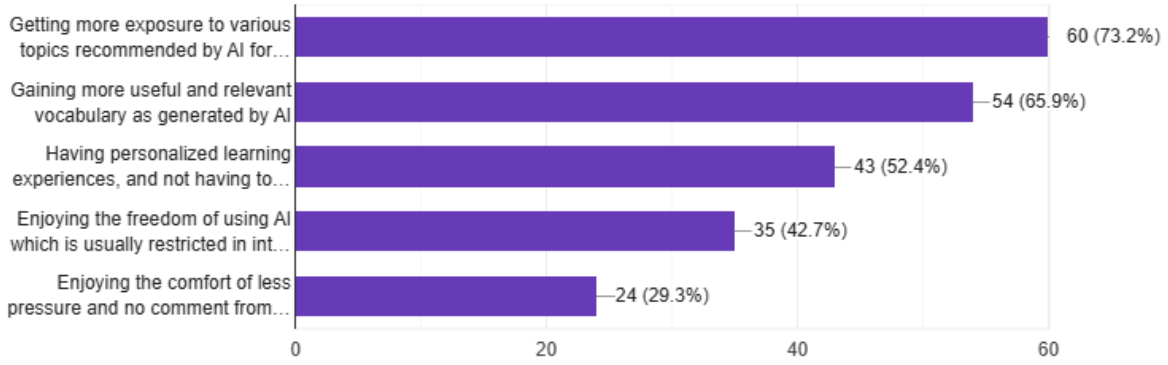
Regarding quantitative analysis, the data reliability can be reinforced (as shown in Table 1): for Q2, a standard deviation of 0.744 is relatively low for a 5-point scale, showing a strong consensus among the cohort. Most responses are around the *"Useful"* and *"Very useful"* categories. A mean score of 4.20 shows that the average respondent finds using AI tools to facilitate their interpreting training to be *"Useful"*.

4.2 Addressing Research Question 2

Student Perceived Benefits in Using AI to Support Self-Practice in Interpretation Courses

Figure 3

Common perceived benefits of utilizing AI tools in self-practice of interpreting skills



When being asked about common benefits of AI support in facilitating self-practice, many students admitted such benefits as *"gaining more useful and relevant vocabulary as generated by AI"*, accounting for almost 66%. of the students. *"Having personalized learning experiences, and not having to rely on peers for practice"* is noted by more than half of the students under survey. Many students pointed to the activity's benefits such as *"enjoying the freedom of using AI which is usually restricted in interpreting courses"* and *"enjoying the comfort of less pressure and no comment from instructor/peers"* with the proportion of

approximately 43% and 30% respectively. Notably, more than 73% agreed on *"Getting more exposure to various topics recommended by AI for practice"*. Some students put forth their comments:

"Use AI as a real-time partner to simulate live interpreting scenarios and provide immediate feedback on linguistic nuances and delivery speed."

"AI can improve interpreting through real-time practice, instant feedback, and personalized learning."

"Using ChatGPT helps me practice interpreting more easily by generating scripts and giving quick feedback."

Similar comments on the benefits were also noted from the focus group interviews:

"AI helps me practice interpreting with instant feedback and flexible topics." (S4)

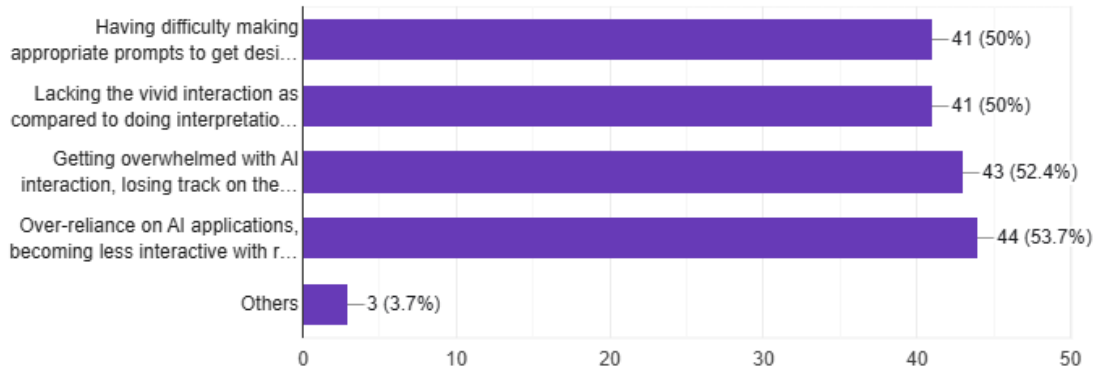
"AI tools like ChatGPT, Gemini help me practice interpreting by providing many topics and speeches. I can practice anytime, which improves my vocabulary and understanding." (S6)

4.3 Addressing Research Question 3

Common Challenges for Students in Utilizing AI Support for Interpreting Training, and Student Recommendations for Improvement

Figure 4

Common disadvantages/difficulties in employing AI tools to facilitate interpreting practice



The graph shows the challenges of using AI as a facilitator for interpreting training are recorded not critical. Regarding the difficulties/disadvantages in utilizing AI tools for interpreting training, *"having difficulty making appropriate prompts to get desired AI-generated content for practice"* represents 50%. Similarly, also 50% of the students admitted *"Lacking the vivid interaction as compared to doing interpretation with people"* as one of the disadvantages. More than half of the students agreed on *"Getting overwhelmed with AI interaction, losing track on the major task_practicing interpreting"* (52%). And the most notable challenge facing the students is *'Over-reliance on AI applications, becoming less interactive with realistic situations for interpretation",* which was confirmed by more than 54%. This figure echoes the observation-generated fact showing a phenomenon that while students were engaged in interacting with AI-mediated tools, many apparently allocated more time for other related tasks than for performing their interpreting job. Some of students' comments from the survey:

"Users should use it actively, not depend on it."

"It cannot fully simulate real-life pressure or natural speaking situations."

"AI speech is sometimes not natural…I should also review my performance to find mistakes. AI is useful, but it should be used in a smart and active way."

Likewise, group interviews echoed similar findings from the survey results. Some students voiced their concerns over common challenges, of which some notable drawbacks can be the lack of vivid interaction as compared to human interactions, and the risk of overwhelm with or over-reliance on employing AI tools.

Students' Recommendations for Improvements

It is noted that many of the students' suggestions and recommendations are insightful, showing their serious attitude in AI employment and reflection. Multiple interesting opinions from students being gained from their practice using the approach were notably recorded. Their inputs centered around proposing some strategies for utilizing AI-mediated tools for interpreting training effectively and recommending some useful AI tools as facilitators in self-practice.

Some note-worthy features in students' recommendation as below:

"I think the Natural Reader is a really useful tool for interpretation skills."

"Elevenlabs"

"Use Gemini, Chatgpt, Copilot"

"Increase difficulty gradually and practice both languages direction, focusing on specialized topics."

"Verify the information provided by the AI because it may not be accurate."

"Use AI as a real-time partner to simulate live interpreting scenarios and provide immediate feedback on linguistic nuances and delivery speed."

" To improve efficiency, I should combine AI with self-recording, practice under time pressure, and balance AI use with real interaction practice."

Similar suggestions and recommendations were echoed in comments given by participants in the focus group interviews:

"...The most efficient use of AI is to make practice more deliberate. Many learners use AI only to generate random texts or check vocabulary. That is useful, but limited. A better approach is to use AI at each stage of the interpreting cycle: input, performance, feedback, and reflection (S1)

"I recommend use Grok because this AI platform can create some ideas"(S6)

"... To improve, I should use different tools together, practice with increasing difficulty, and simulate real situations. I should also review my performance to find mistakes. AI is useful, but it should be used in a smart and active way."(S8)

"Balance between using AI for interpreting and studying."(S5)

" To improve, I should simulate real situations, review my performance, and combine AI with real-life practice."(S9)

5. Discussion

5.1 Discussion of Findings

The empirical findings provided valuable insights into one of the most significant contemporary issues in ELT: integrating Artificial Intelligence (AI) into pedagogical practice. Despite the advantages and disadvantages associated with the use of AI-powered platforms in interpreting training, AI integration has been apparently welcomed by both teachers and students as supplementary facilitators for teaching and learning interpretation. This approach aligns with the necessity of pedagogical adaption in the era of AI being highlighted by Cui et al. (2025) and Šveda & Poláček (2025).

Regarding the benefits for students in self-practice, AI-powered platforms seemingly outperform many existing technological applications in several aspects. According to students' feedback, AI has offered tremendous facilitation such as providing adaptive feedback and assistance 24/7, thereby enhancing self-regulated learning and autonomy. These findings are consistent with Consuelo Sáiz-Manzanares et al. (2023) and Ng, Tan, and Leung (2024) claiming that AI-powered platforms can help improve learners' autonomy. Data collected from group interviews also revealed another advantage: by interacting with AI-powered platforms and receiving instant feedback, students has immensely improved their cognitive process reflex, which was pointed out by Chien et al. (2024) and Yu et al. (2025). Last but not least, increased motivation was also reported by students due to the freedom of utilizing AI support and the comfort of learning in a more relaxing and low-stress practice environment. This prominent feature aligns with findings from Tran (2026) pointing out that the AI tools can be integrated to positively affect learners' motivation, together with reduced anxiety and increased confidence.

Regarding the obstacles encountered learners while using AI to facilitate their interpreting training, the research indicates several adverse implications. Students may find it difficult to generate prompts in utilizing AI platforms; to mitigate this disadvantage, it is advocated that explicit guidelines should be provided. Another drawback in practicing with AI-powered platforms is the feeling of missing human vivid interaction. Students reported the lack of the nuanced human cognition in complex, dynamic communicative contexts. This reflection aligns with the concerns from literature review by Tang, L. (2025) and Gómez & Pinazo (2026) arguing about AI's incapability in replacing human interaction. In addition, a primary challenge involves embedding AI into self-practice without compromising learners' active cognitive involvement and skill acquisition. In other words, learners' efficient autonomy can be questioned because their engagement in self-practice is predominantly self-regulated and lacks formal oversight. The matters of overwhelm or over-reliance with AI interaction is disputable with detrimental effect of degrading learners' critical thinking and analytical reasoning. This challenge echoes the findings from Darvishi et al. (2024) who also warned that student agency can be negatively impacted if over-reliance occurs. In interpreting training, over-reliance or misuse of AI tools may result in even more deteriorating consequence. In light of literature review, over-reliance on generative AI supposedly degrades the cognitive working memory reflexes required in real-time interpretation. Since interpreting skills require instant reflexes, it takes prompt and effective cognitive process that can be achieved only by consistent practice on one's own. With AI over-reliance, students tend to utilize AI not only as a facilitator but also as a substitute interpreter for them, which may deprive them of the chance to train their cognitive process reflexes, resulting in poor skills performance. Therefore, as mentioned previously, AI integration in interpreting training aims to facilitate learners with their self-practice as supportive tools only. Importantly, AI must not replace students in performing the major function - interpreting if students are to improve their skills.

It is assumed that certain challenges are inevitable in employing advanced technology. In practice, considerable attention has been given to the underlying challenges associated with AI integration into interpreting training. Overall, based on the literature review and empirical findings of this study, the benefits of AI integration in interpreting training, particularly for learners' self-practice, supposedly outweigh the limitations in various respects. Benefits should be optimized while challenges can be mitigated with consideration to students' reflections and recommendations.

Regardless of whether this practice is fully supported or debated, the integration of AI remains an contemporary trend in English Language Teaching (ELT) in general and interpreting training in particular. Moreover, this pedagogical approach is consistent with the learner-centered philosophy of education which emphasizes the promotion of learner autonomy and the maximization of learners' benefits.

5.2. Limitation of The Study

As common sense, certain limitations of this study are inevitable. Firstly, the research was conducted on a relatively small number of students at the Faculty of Foreign Languages of HUFLIT. For more empirical insights and profound pedagogical implications for AI integration in this particular context, more interpretation classes should be observed with more instructors and students involved in different academic settings. Secondly, further research should also cater to keeping track of students' progress over a given period of implementation for proven efficiency of the pedagogical approach. Lastly, these advanced digital applications have recently been integrated into pedagogical practices for several years only, it takes more time for proven efficacy, including scrutinizing and evaluating the integrity of AI feedback to students' interpreting outputs, and investigating negative aspects of the approach such as misuse and over-reliance.

5.3. Recommendations

With a view to enhancing the benefits and alleviating the drawbacks of AI integration into interpreting training, following strategies are recommended. First, students should be made fully aware of the intent of utilizing AI in their self-practice where interpreting skills acquisition fundamentally necessitates rigorous, independent practice. The actual cognitive task of interpretation remains strictly students' responsibility - a boundary that must be maintained for genuine skills improvement. Accordingly, the extent and restrictions of using AI in self-practice should be clarified to students at the outset: AI integration must comply with the provided guidelines; AI serves as a mediator/ facilitator only and will never replace students in producing interpreting outputs at any cost if students are to hone their skills. Second, students should be warned against the adverse effects of AI over-reliance on their cognitive competence. Thus, they should stay alert, using critical thinking and logical reasoning to determine the effectiveness of AI feedback, not be overly optimistic or complacent with AI comments that can be flattering occasionally. To avoid inflated expectations, continuous effort is essential. On the part of instructors, this approach advocates for transparent, explicit guidelines for proper utilization over prohibition that may lead to misuse. No ban is imposed on student use of AI-powered platforms, but ethical use should be taught. Furthermore, oral tests should be given in class on a regular basis to monitor and evaluate students' genuine performance of interpreting skills, keeping track of their continued progress. Last but not least, students' performance from self-study with AI assistance should be considered for on-going assessment, not for formal testing, which should be well informed to students in advance in order to prevent intentional cheating or misuse.

6. Conclusion

To conclude, in tertiary education, AI integration has recently transcended a mere trend to become an essential pedagogical requirement. In interpreting training where AI applications were initially restricted, AI-powered platforms have become gradually recognized as beneficial and supportive. However, the extent and methodology of AI integration in interpreting training require careful consideration by educational stakeholders. In alignment with the literature review, this study has yielded empirical findings with insightful implications regarding enhancing student engagement through AI-assisted self-practice in interpretation courses. Despite certain inevitable drawbacks, the benefits of utilizing AI as a facilitator for self-practice in interpreting training have prevailed in many respects. AI integration makes self-regulated learning process more personalized and engaging for students. Particularly, AI integration for self-practice in interpreting training may help enhance students' engagement and promote their autonomy, aligning with the learner-centered pedagogical philosophy that should be embraced.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Chien, C., H. Chan, and H. Hou. 2024. "Learning by Playing with Generative AI: Design and Evaluation of a Role-Playing Educational Game with Generative AI as Scaffolding for Instant Feedback Interaction." Journal of Research on Technology in Education 57 (4): 894–913. 1–20. https://doi.org/10.1080/15391523.2024.2338085.

Consuelo Sáiz-Manzanares, M., R. Marticorena-Sánchez, L. J. Martín-Antón, I. González Díez, and L. Almeida. 2023. "Perceived Satisfaction of University Students with the Use of Chatbots as a Tool for Self-Regulated Learning." Heliyon 9 (1): e12843. 10.1016/j.heliyon.2023.e12843.

Cui, F., Li, D., & Zhuang, C. (2025). Introduction: transforming translation education through artificial intelligence. The Interpreter and Translator Trainer, 19(3-4), 227-233. https://doi.org/10.1080/1750399x.2025.2561258

Darvishi, A., H. Khosravi, S. Sadiq, D. Gašević, and G. Siemens. 2024. "Impact of AI Assistance on Student Agency." Computers and Education 210:104967. https://doi.org/10.1016/j.compedu.2023.104967.

Giustini, D., & Dastyar, V. (2024). Critical AI literacy for interpreting in the age of AI. Interpreting and Society, 4(2), 196-213. https://doi.org/10.1177/27523810241247259

Gómez, L. P., & Pinazo, E. P. (2026). Beyond the hype: why AI still cannot replace human translators and interpreters in high-stakes contexts. FITISPos International Journal, 13(1), 13-24. https://doi.org/10.37536/fitispos-ij.2026.13.1.450

Gutiérrez, C. C. (2025). AI-assisted interpreting: Valuable tool for professional interpreters or job displacement? Traduction et Langues, 24(1), 263–281.

Hidayat, D. N., Ramadhan, S., Mahlil, M., Mason, J., Hartono, R., & Muslimin, A. I. (2025). Research trends in artificial intelligence in english language teaching (ELT): a bibliometric analysis. Studies in English Language and Education, 12(2), 560-578. https://doi.org/10.24815/siele.v12i2.39339

Ng, D. T. K., C. W. Tan, and J. K. L. Leung. 2024. "Empowering Student Self-Regulated Learning and Science Education Through chatgpt: A Pioneering Pilot Study." British Journal of Educational Technology 55 (4):1–26. https://doi.org/10.1111/bjet.13454.

Pöchhacker, F., & Liu, M. (2024). Interpreting technologized. Interpreting. International Journal of Research and Practice in Interpreting, 26(2), 157-177. https://doi.org/10.1075/intp.00112.poc

Šveda, P., & Poláček, I. (2025). Interpreter training in the age of AI. L10N Journal.

Tang, L. (2025). The lack of other minds as the lack of coherence in human–AI interactions. Philosophies, 10(4), 77. https://doi.org/10.3390/philosophies10040077

Tang, J. (2023). Artificial intelligence-based needs analysis for English specific purposes in digital environment. Learning and Motivation, 83, 101914. https://doi.org/10.1016/j.lmot.2023.101914

Tran, T. M. P. (2026). Current use of AI translation and interpreting tools among Vietnamese students: A preliminary survey of their impact on motivation to learn Japanese. VNU

Journal of Foreign Studies, 42(2), 127-143. https://doi.org/10.63023/2525-2445/jfs.ulis.5634

Yu, Y., Wei, W., & Chen, Z. (2025). Comparing learners' engagement strategies with feedback from a generative AI chatbot and peers in an interpreter training programme: a quasi-experimental study. The Interpreter and Translator Trainer. https://doi.org/10.1080/1750399x.2025.2533069

Ethics Statement

The research was conducted in accordance with accepted ethical standards for academic research, including voluntary participation, informed consent, and protection of participant confidentiality.

Bionote

Dang Thi Anh Tuyet is a Master of TESOL and a Bachelor of Economics with almost 20 years of teaching English in both academic and business settings. Tuyet is currently a lecturer at Faculty of Foreign Languages, HUFLIT, a dynamic working environment. She is a devoted lecturer, passionate about bringing values to learners, connecting theory with the reality.

Appendices

Appendix A. Classroom Observation Checklist

Observation Criteria	High	Moderate	Low	Notes
1. Students show interest in employing AI tools such as ChatGPT, Gemini, etc. in interpreting self-practice.	☐	☐	☐	
2. Students exhibit engagement, actively interacting with AI as a facilitator in interpreting self-practice.	☐	☐	☐	
3. During implementing, students mostly follow the guidelines, only using AI as a facilitator for self-practice.	☐	☐	☐	
4. Students have difficulty in making appropriate prompts for AI chatbots.	☐	☐	☐	
5. Students exhibit overwhelm with AI interaction, not focusing on training interpreting skills.	☐	☐	☐	

Appendix B. Survey Questionnaire

Survey on Utilizing AI support for Interpreting Practice

1. How interesting do you find using AI tools to facilitate your practice of interpreting skills?

B I U

- ○ Very interesting
- ○ Interesting
- ○ Neutral
- ○ Boring
- ○ Very boring

2. How useful do you find using AI tools to facilitate your practice of interpreting skills?

B I U

- ○ Very useful
- ○ Useful
- ○ Neutral
- ○ a little useful
- ○ Not useful at all

3. In your opinion, what are common benefits of using AI tools to facilitate your practice of interpreting skills? (more than one choice is possible)

- [ ] Getting more exposure to various topics recommended by AI for practice
- [ ] Gaining more useful and relevant vocabulary as generated by AI
- [ ] Having personalized learning experiences, and not having to rely on peers for practice
- [ ] Enjoying the freedom of using AI which is usually restricted in interpreting courses
- [ ] Enjoying the comfort of less pressure and no comment from instructors/peers

4. What are some disadvantages/difficulties you face when using AI tools to facilitate your practice of interpreting skills? (more than one choice is possible)

- [ ] Having difficulty making appropriate prompts to get desired AI-generated content for practice
- [ ] Lacking the vivid interaction as compared to doing interpretation with people
- [ ] Getting overwhelmed with AI interaction, losing track on the major task_practicing interpreting
- [ ] Over-reliance on AI applications, becoming less interactive with realistic situations for interpretation
- [ ] Others

5. Your comments and recommendations for improving the efficiency of using AI to facilitate your practice of interpreting skills in general.

Appendix C. Interview Questions

1. In your opinion, what are some common benefits of using AI tools to facilitate your self-practice of interpreting skills?

2. What are some disadvantages/difficulties you face when using AI tools to facilitate your self-practice of interpreting skills?

3. Your comments and recommendations for improving the efficiency of using AI to facilitate your self-practice of interpreting skills in general.

#186 ARTIFICIAL INTELLIGENCE IN ENHANCING LISTENING AND NOTE-TAKING SKILLS: EVIDENCE FROM FISRT-YEAR STUDENTS IN VIETNAM

Tran Tan Phat[1], Le Ngoc Anh Thu[2]

Abstract

The rapid integration of Artificial Intelligence (AI) in higher education has created new opportunities to enhance language learning, particularly in listening comprehension and note-taking skills. This research investigates the effectiveness of AI-assisted instruction in improving listening and note-taking abilities among first-year English as a Foreign Language (EFL) students at a private university in Vietnam (HUFLIT). Adopting a mixed-methods design, the research involved 80 participants divided into an experimental group, which engaged with AI-based tools for listening practice and structured note-taking, and a control group receiving conventional instruction. Quantitative data were collected through pre- and post-tests and analyzed using paired and independent samples t-tests, while qualitative insights were obtained from student questionnaires and reflective journals. The findings indicate that students in the experimental group demonstrated statistically significant improvements in both listening comprehension and note-taking performance compared to their counterparts in the control group. Moreover, participants reported increased engagement, autonomy, and confidence when using AI tools, although some challenges related to over-reliance and technical limitations were noted. The research underscores the pedagogical potential of AI in supporting skill development in EFL contexts and highlights the need for guided implementation to maximize learning outcomes. These results contribute to the growing body of research on AI in language education and offer practical implications for curriculum design and instructional practices in similar tertiary settings.

Key words: Artificial Intelligence, Listening skills, Note-taking, EFL students, First-year student

1. Introduction

The rapid integration of Artificial Intelligence (AI) into higher education has fundamentally transformed pedagogical practices, particularly in English as a Foreign Language (EFL) contexts. At Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), a private institution in Vietnam with a specialized focus on foreign languages and information technology, AI tools have popularly used language learning environments. These tools range from AI-powered chatbots and grammar checkers like ChatGPT and Grammarly to speech recognition applications such as ELSA Speak and

[1]*Faculty of Foreign Languages, Ho Chi Minh University of Banking*
Cellphone: +84903821443; Email: phattt@hub.edu.vn
[2]*Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology*

Google Translate. The university has demonstrated a commitment to educational technology, with ongoing research examining students' AI competency and the effectiveness of AI-assisted instruction in skill development (Nguyen et al., 2026; Tran et al., 2025). This technological shift presents both opportunities and challenges for first-year students who are navigating the transition from secondary to tertiary education while simultaneously adapting to AI-enhanced learning paradigms.

First-year students at HUFLIT exhibit a relationship with AI in developing English listening and note-taking skills. Research conducted at the university indicates that while students self-report high levels of AI competency, there exists a notable discrepancy between perceived and actual understanding of AI tools in language learning. Furthermore, comparative studies between English majors and non-English majors at HUFLIT have revealed that first-year students demonstrate substantial readiness for AI integration, characterized by strong perceptions of usefulness and ease of use. However, findings highlight concerns regarding over-reliance on AI tools, accuracy issues, and the need for institutional guidance. These findings suggest that while first-year students are enthusiastic about AI adoption, their effective utilization of these tools for specific skills like listening comprehension and note-taking remains underexplored. Recent research on AI-assisted note-taking has demonstrated potential benefits, including reduced cognitive load and enhanced lecture comprehension, as AI tools allow students to focus on content rather than transcription. However, the application of such tools to listening texts in the Vietnamese EFL context requires systematic investigation (Hien, 2025; Phan, 2023). Therefore, this research addresses the following research questions:

1. What are HUFLIT first-year students' perceptions of using AI in learning English?
2. How do AI tools impact HUFLIT first-year students' English learning skills?
3. How did they apply note-taking skills in listening texts?

2. Literature Review

2.1. The Research in Academic Listening and Note-Taking

Academic listening represents a uniquely demanding task for English as a Foreign Language (EFL) learners, requiring simultaneous processing of oral input, comprehension of content, and generation of written records. Note-taking while listening imposes substantial cognitive load, as learners must allocate limited attentional resources between understanding incoming speech and capturing key information. This dual-task challenge is particularly acute for first-year university students, who often enter tertiary education with minimal prior training in systematic note-taking strategies (Shizuka & Ryan, 2026; Tran et al., 2025).

Recent research has substantially advanced understanding of what constitutes effective note-taking. A significant study by Shizuka & Ryan (2026) involving 195 Japanese EFL first-year students demonstrated that note quality – specifically organization and inclusion of main ideas and supporting details – correlates more strongly with listening comprehension outcomes than the mere quantity of words produced. Their findings revealed that well-organized notes, characterized by clear hierarchical structures and appropriate use of abbreviations and symbols, serve as stronger predictors of performance on standardized listening assessments such as the TOEFL ITP. These results challenge earlier research that relied primarily on word count as a metric of note-taking proficiency (Siegel, 2022).

Furthermore, empirical evidence indicates that note-taking skills are teachable. Shizuka & Ryan (2026) documented significant improvements in both note quality and listening comprehension following structured strategy instruction. The encoding function of note-taking

– the cognitive processing that occurs during the act of writing – appears to enhance retention and comprehension independent of the review function. However, research by Kim (2014) in the Korean EFL context revealed a notable discrepancy: while students express positive perceptions of note-taking importance, their actual implementation remains limited without explicit training. This gap between perception and practice underscores the necessity of pedagogical intervention.

2.2. The Influence of Artificial Intelligence on Education

Artificial Intelligence has emerged as a transformative force in educational contexts, with systematic reviews demonstrating its capacity to enable personalized learning pathways, improve instructional practices, and enhance feedback mechanisms. A systematic literature review by Adil (2025) synthesizing 39 peer-reviewed studies from 2020 to 2024 identified both benefits and persistent challenges, including infrastructural limitations, inadequate teacher training, and ethical concerns regarding algorithmic bias and data privacy. Notably, the review highlighted a shortage of classroom-based empirical evidence, suggesting that while AI's potential is widely acknowledged, rigorous investigation of its pedagogical effectiveness remains incomplete.

In language education specifically, AI tools have proliferated rapidly. Speech-to-text (STT) technologies provide immediate transcripts that facilitate noticing of pronunciation errors, while text-to-speech (TTS) applications offer adjustable playback speeds and customizable voices that accommodate varying proficiency levels (Jongbum, 2025). More recently, generative AI platforms such as NoteBookLM have enabled transformation of written texts into podcast-style dialogues, restructuring input genre to potentially reduce cognitive load and enhance comprehension. Liu and Zhao (2025) conducted a quasi-experimental study demonstrating that EFL learners exposed to NoteBookLM-generated podcasts achieved significantly higher listening comprehension scores and reported greater reductions in listening anxiety compared to control groups receiving traditional news recordings, while holding content constant.

2.3. Students' Perceptions of AI in Listening Classes

Understanding learner perceptions of AI tools is critical, as technology acceptance fundamentally shapes engagement and outcomes. Research by Jongbum (2025) with 114 Korean university EFL students revealed generally positive attitudes toward AI-generated listening materials, with students particularly valuing adjustable audio playback speed and the option to revisit content – features that foster autonomous learning. Students acknowledged that repeated exposure to structured passages, dialogs, and summaries reinforced vocabulary acquisition and supported comprehension development (Irgin, 2025).

However, substantial challenges were identified. Jongbum's participants reported unnatural intonation patterns, robotic voice delivery, and pronunciation inaccuracies in AI-generated speech, issues that were especially problematic for lower-proficiency learners who rely more heavily on prosodic cues for comprehension. These findings align with Meniado (2023) and Ali (2025), who documented concerns regarding unnatural prosody and limited contextual nuance in synthetic speech. Collectively, the literature suggests that a hybrid approach – combining AI's adaptability with instructor-led refinement – optimizes both skill development and positive affective responses.

2.4. Acquiring Listening Comprehension with AI-Enhanced Note-Taking

Complementing these findings, research from the Diplomatic Academy of Vietnam by Do & Dang (2025) explored AI-generated speech simulations to support short-term memory and consecutive interpreting training. While students using AI speech demonstrated greater progress on tasks requiring higher cognitive effort, they continued to struggle with note-taking and vocabulary, highlighting the persistent need for instructor guidance even within AI-enhanced environments. The theoretical framework underlying AI-enhanced note-taking draws upon Mayer's (1996) SOI model (Selection, Organization, Integration), which posits that effective learning requires learners to select relevant information, organize it coherently, and integrate it with prior knowledge. AI tools may scaffold these processes by providing real-time transcription, highlighting key concepts, or suggesting organizational structures, thereby reducing cognitive load and enabling deeper processing (Al-Ghazo, 2023; Siegel, 2020; Shizuka & Ryan, 2026).

3. Methodology

3.1. Research Design

This study adopted a mixed-methods sequential explanatory design, which integrates quantitative and qualitative approaches to provide a comprehensive understanding of the research problem (Creswell & Poth, 2018; Dörnyei, 2007). Mixed-methods research is particularly suited for educational investigations, as it enables the researcher to capture both statistical generalizability and contextual depth – strengths that compensate for the limitations of each individual approach when used in isolation. As Creswell (2014) notes, mixed-methods designs mirror everyday problem-solving practices and are essential when research questions cannot be adequately addressed by a single data type. The quantitative strand employed a quasi-experimental pretest-posttest control group design, which is widely recognized as appropriate for evaluating pedagogical interventions in classroom settings. Eighty first-year students were assigned to either an experimental group (receiving AI-assisted instruction) or a control group (receiving conventional instruction). The qualitative strand utilized semi-structured reflective journals and open-ended questionnaire items to capture participants' perceptions, challenges, and strategic approaches to AI-enhanced note-taking. This sequential design allowed quantitative findings on skill improvement to be subsequently explained and enriched by qualitative insights into learners' subjective experiences.

3.2. Research Participants

The research was conducted at the FFL, HUFLIT. A total of 80 first-year EFL students participated in the research. Participants were drawn from two listening classes to preserve the natural classroom ecology, a practice consistent with quasi-experimental educational research. The sample size aligns with comparable studies in the Vietnamese EFL context, where similar investigations have employed 80 to 100 participants for mixed-methods designs.

All participants were first-semester students with A2 to B1 CEFR proficiency levels, representing the typical English competency range of newly enrolled students at Vietnamese private universities. Students were not randomly assigned to conditions due to institutional scheduling constraints; instead, one intact class was designated as the experimental group (n=40) and the other as the control group (n=40). This non-random assignment is acknowledged as a limitation of the quasi-experimental design, yet it reflects authentic classroom conditions and enhances external validity.

3.3. Research Instruments

Four instruments were employed to ensure triangulation of data sources:

Listening Comprehension and Note-Taking Tests (Pre- and Post-): Both tests were developed based on the TOEFL ITP listening section format, which has demonstrated reliability in prior EFL listening research. The pre-test assessed baseline listening comprehension and note-taking ability using a 30-minute academic lecture excerpt. The post-test, equivalent in difficulty and structure, measured gains following the eight-week intervention. Note-taking performance was evaluated using a rubric adapted from Shizuka & Ryan (2026), assessing organization, inclusion of main ideas, supporting details, and use of abbreviations/symbols.

AI-Enhanced Note-Taking Intervention Materials: For the experimental group, a suite of AI tools was integrated into listening instruction. These included STT applications (Google Speech-to-Text and VALSEA Realtime ASR) providing live transcripts, AI summarization tools (ChatGPT) generating structured lecture notes, and AI-based clarification tools (VALSEA Clarifications) offering simplified Vietnamese explanations of English lecture content. These tools were selected based on their demonstrated efficacy in Southeast Asian EFL contexts, particularly their capacity to handle Vietnamese-accented English and code-switching between Vietnamese and English – features critical to HUFLIT's instructional environment.

Student Perception Questionnaire: A 20-item Likert-scale questionnaire (1 = strongly disagree to 5 = strongly agree) was administered post-intervention to the experimental group. Items measured perceived usefulness, ease of use, engagement, and challenges associated with AI tools. The questionnaire was adapted from validated instruments used in previous Vietnamese EFL technology studies (Luu & Doan, 2025), with Cronbach's alpha exceeding .85 in pilot testing.

Reflective Journals: Experimental group participants submitted weekly reflective journals documenting their note-taking strategies, perceptions of AI tools, and difficulties encountered. Journal prompts guided students to describe specific instances of AI use and evaluate its impact on their listening comprehension.

3.4. Data Collection

Data collection occurred over eight weeks during the first semester of the 2025 – 2026 academic year, following a structured timeline:

Week 1 (Pretest): Both groups completed the listening comprehension pretest under standardized conditions. Participants listened to a 30-minute academic lecture while taking notes, followed by 20 multiple-choice comprehension questions. Notes were collected for subsequent analysis.

Weeks 2–7 (Intervention): The control group received conventional listening instruction: lecture listening, post-listening comprehension questions, and teacher-modeled note-taking strategies. The experimental group received AI-enhanced instruction: (a) live transcription via STT tools during initial listening; (b) guided comparison of personal notes with AI-generated transcripts; (c) use of ChatGPT for summary generation and vocabulary clarification; and (d) structured practice integrating AI outputs into conventional note-taking frameworks. Each 90-minute session concluded with reflective journaling.

Week 8 (Posttest and Questionnaire): Both groups completed the listening posttest under identical conditions. The experimental group additionally completed the perception questionnaire and submitted their reflective journal compilations.

4. Results and Discussion

4.1. Analysis of Survey Results from 80 First-Year Students

Quantitative analysis revealed statistically significant between-group differences. The experimental group reported substantially higher mean scores across all perception dimensions (M=4.12, SD=0.58) compared to the control group's perceptions of conventional methods (M=3.24, SD=0.71), t(78)=6.14, p<.001. Notably, 87.5% of experimental participants agreed or strongly agreed that AI tools improved their listening comprehension, while 82.5% reported reduced anxiety during listening tasks. These findings align with Luu and Doan (2025), who documented similar positive perceptions among Vietnamese EFL learners using ChatGPT for listening practice.

However, challenges were identified. Approximately 45% of experimental participants expressed concerns regarding AI-generated speech naturalness and transcription accuracy, particularly with Vietnamese-accented English and specialized vocabulary. These concerns mirror findings from Jongbum (2025) in the Korean EFL context and Meniado (2023) regarding prosodic limitations of synthetic speech.

Table 1

Descriptive Statistics of Student Perceptions (Post-Intervention)

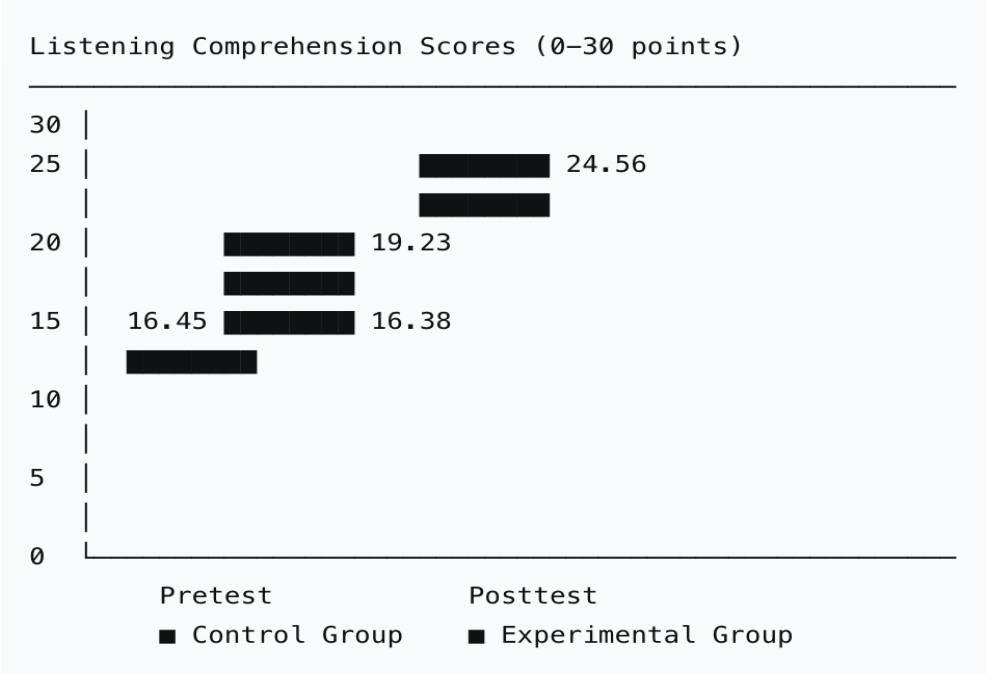
Perception Dimension	Experimental Group (n = 40) Mean (SD)	Control Group (n = 40) Mean (SD)	t-value	p-value
Perceived Usefulness	4.28 (0.52)	3.15 (0.68)	8.34	<.001
Ease of Use	4.05 (0.61)	3.30 (0.72)	5.12	<.001
Engagement	4.18 (0.55)	3.28 (0.69)	6.45	<.001
Reduced Listening Anxiety	3.98 (0.70)	3.22 (0.80)	4.56	<.001
Overall Perception	4.12 (0.58)	3.24 (0.71)	6.14	<.001

Table 1 reveals that the experimental group, which received AI-enhanced instruction, reported significantly higher perceptions across all measured dimensions compared to the control group. The largest difference emerged in perceived usefulness (4.28 vs. 3.15, t=8.34), indicating that students who actively used AI tools recognized their value for listening comprehension. Reduced listening anxiety showed the smallest but still significant difference (3.98 vs. 3.22), suggesting that while AI helps alleviate some stress associated with real-time listening, anxiety is not entirely eliminated. These findings align with Luu and Doan (2025),

who reported that Vietnamese EFL learners experience moderate anxiety reduction when using ChatGPT for listening tasks. The consistently higher standard deviations in the control group (e.g., 0.71 vs. 0.58 for overall perception) suggest greater variability in attitudes toward conventional instruction, whereas AI-enhanced instruction produced more uniformly positive perceptions.

Figure 1

Pre- and Post-Intervention Listening Comprehension Scores



The bar chart illustrates a striking differential improvement between groups. At pretest, both groups demonstrated nearly identical means (16.45 vs. 16.38), confirming successful randomization equivalence. Following the eight-week intervention, the experimental group's mean score increased to 24.56 – a gain of 8.18 points (50% improvement) – while the control group rose to only 19.23 – a gain of 2.78 points (17% improvement). The gap between groups widened from 0.07 points at pretest to 5.33 points at posttest. This visually reinforces the large effect size (Cohen's d=2.97 for experimental vs. 0.94 for control) reported previously. Such dramatic improvement suggests that AI tools – particularly real-time transcription and AI-generated summaries – enabled students to overcome the dual-task burden of simultaneous listening and note-taking, consistent with Mayer's (1996) SOI model where AI scaffolds the selection and organization phases.

Table 2

Note-Taking Quality Scores by Assessment and Group

Assessment Point	Control Group (n=40) Mean (SD)	Experimental Group (n=40) Mean (SD)	Mean Difference	% Improvement
Pretest (Week 1)	5.28 (1.38)	5.34 (1.42)	+0.06	—

Assessment Point	Control Group (n=40) Mean (SD)	Experimental Group (n=40) Mean (SD)	Mean Difference	% Improvement
Posttest (Week 8)	6.45 (1.67)	8.67 (1.15)	+2.22	—
Gain Score	+1.17	+3.33	+2.16	—
% Improvement	22.2%	62.4%	—	40.2% higher

Table 2 demonstrates that both groups improved in note-taking quality, but the experimental group's gain (3.33 points) nearly tripled that of the control group (1.17 points). The 62.4% improvement in the experimental group, compared to only 22.2% in the control group, indicates that AI tools substantially enhanced students' ability to produce organized, content-rich notes. Notably, the experimental group's posttest standard deviation (1.15) was smaller than its pretest standard deviation (1.42), suggesting that AI tools not only raised average performance but also reduced variability among learners – a crucial finding for educators concerned about unequal outcomes. Conversely, the control group's standard deviation increased from 1.38 to 1.67, indicating that conventional instruction may widen performance gaps between stronger and weaker note-takers. This pattern supports Jongbum (2025), who found that AI tools particularly benefit lower-proficiency learners by providing structured scaffolding.

4.2. Three Quantitative and Qualitative Models

Model 1: Perceived Usefulness-Actual Usage Correlation Model

Quantitative analysis using Pearson correlation demonstrated a moderate positive relationship between perceived usefulness of AI tools and frequency of tool usage ($r=0.62$, $p<.01$). However, a notable discrepancy emerged: while 92.5% of students perceived AI as useful for listening, actual usage patterns showed that only 67.5% regularly employed AI tools beyond initial training sessions. Qualitative journal entries explained this gap: *"AI transcripts help me catch missed words, but sometimes I spend too much time comparing rather than listening"*.

Model 2: Note-Taking Quality Improvement Model

Pre- and post-intervention note-taking assessments, evaluated using Jin and Webb's (2024) rubric (0-12 points scale), revealed differential improvement patterns. The experimental group's mean score increased from 5.34 (SD=1.42) to 8.67 (SD=1.15), representing a 62.4% improvement, compared to the control group's increase from 5.28 (SD=1.38) to 6.45 (SD=1.67), a 22.2% improvement. Independent t-test confirmed statistical significance, $t(78)=6.89$, $p<.001$. Qualitative analysis identified three distinct note-taking approaches adopted by experimental participants: AI-assisted linear notes (42.5%), where students used AI transcripts as base texts for annotation; comparative dual-column notes (35%), where personal notes were written alongside AI-generated summaries for verification; and AI-independent strategic notes (22.5%), where students deliberately avoided AI to practice unaided listening.

Model 3: Listening Comprehension Gain Score Model

Pre-test & post-test listening comprehension scores (TOEFL ITP format, 0-30 points) showed significant gains for both groups, but with substantially larger effect size in the

experimental group. Control group mean scores increased from 16.45 (SD=2.89) to 19.23 (SD=2.97), Cohen's d=0.94. Experimental group mean scores increased from 16.38 (SD=2.76) to 24.56 (SD=2.54), Cohen's d=2.97. The interaction effect (Group × Time) was highly significant, $F(1,78)=34.21$, $p<.001$, $\eta^2=0.31$, indicating a large effect of AI intervention.

Table 3

Correlation Matrix of Key Variables (Experimental Group)

Variable	1	2	3	4	5
1. Perceived Usefulness	1.00				
2. Frequency of AI Tool Use	0.62**	1.00			
3. Note-Taking Quality (Posttest)	0.58**	0.71**	1.00		
4. Listening Comprehension (Posttest)	0.53**	0.67**	0.79**	1.00	
5. Self-Reported Anxiety Reduction	0.48*	0.44*	0.39*	0.52**	1.00

*Note: *p < .05; **p < .01; n=40**

Table 3 provides a correlation matrix revealing the interrelationships among key variables. The strongest correlation emerged between note-taking quality and listening comprehension ($r=0.79$, $p<.01$), confirming the essential role of effective note-taking in achieving listening success – a finding consistent with Jin and Webb's (2024) assertion that note quality predicts comprehension outcomes more strongly than word count. Frequency of AI tool use correlated strongly with both note-taking quality ($r=0.71$) and listening comprehension ($r=0.67$), indicating that active engagement with AI tools directly translates into skill improvement. However, the moderate correlation between perceived usefulness and actual usage ($r=0.62$) suggests a perception-action gap: while most students valued AI, not all translated this perception into consistent practice. Notably, anxiety reduction showed only moderate correlations with other variables (ranging from 0.39 to 0.52), suggesting that while AI alleviates some listening stress, anxiety reduction alone does not guarantee skill acquisition – a nuance that educators should consider when designing interventions.

Table 4

Comparison of Pre- and Post-Intervention Note-Taking Rubric Sub-Scores

Rubric Component (Max Points)	Control Group (n=40)			Experimental Group (n=40)		
	Pretest	Posttest	Gain	Pretest	Posttest	Gain
Organization (3 pts)	1.15	1.48	+0.33	1.18	2.35	+1.17

Rubric Component (Max Points)	Control Group (n=40)	Experimental Group (n=40)			
Main Ideas (3 pts)	1.42	1.75	+0.33	1.45	2.28	+0.83
Supporting Details (3 pts)	1.38	1.62	+0.24	1.40	2.10	+0.70
Abbreviations/ Symbols (3 pts)	1.33	1.60	+0.27	1.31	1.94	+0.63
Total (12 pts)	5.28	6.45	+1.17	5.34	8.67	+3.33

Note: Rubric adapted from Jin & Webb (2024).

Table 4 provides a granular breakdown of note-taking quality across four rubric components. The experimental group demonstrated substantial gains across all components, with organization showing the most dramatic improvement (+1.17 points, a 99% increase from pretest). This suggests that AI-generated transcripts and summaries provided students with clear models of how to hierarchically structure information – main topics followed by supporting details. The control group's gains were modest across all components (ranging from +0.24 to +0.33), indicating that conventional instruction alone does little to improve note-taking structure. Interestingly, abbreviations and symbols showed the smallest gain for both groups (+0.27 for control, +0.63 for experimental), suggesting that this skill – which requires automaticity and practice – develops more slowly than content-based components. This finding implies that explicit instruction on abbreviation systems may be necessary even within AI-enhanced environments. As Participant E31 noted: *"AI gives me full sentences. I still need to learn how to shorten them myself for exams without AI."*

Table 5

Thematic Analysis of Reflective Journals (Experimental Group)

Theme	Frequency	Example Quote	Representative Participant
AI as verification tool	32 (80%)	*"I listen first without AI, then check my notes against the transcript. This shows me what I missed."*	E08
Reduced listening anxiety	28 (70%)	*"Knowing I can see the transcript later makes me less nervous during first listening."*	E22
Concern about over-reliance	26 (65%)	*"Sometimes I just wait for AI instead of trying to listen. That's a bad habit I'm fighting."*	E17

Theme	Frequency	Example Quote	Representative Participant
Improved vocabulary acquisition	24 (60%)	*"AI explains unknown words immediately. I learn more words per lecture now."*	E35
Need for instructor guidance	22 (55%)	*"AI is powerful, but I don't always know how to use it best. Teacher shows us strategies."*	E12
Technical frustration	18 (45%)	*"Slow internet makes AI transcription lag. Very annoying during live listening."*	E03

Note: Multiple themes could be reported per participant; percentages reflect proportion of 40 participants mentioning each theme at least once.

Table 5 summarizes qualitative findings from experimental group participants' weekly reflective journals. The most frequently reported theme (80%) was AI as a verification tool, indicating that students primarily used AI not as a first-line listening aid but as a post-listening check – a metacognitively sophisticated strategy that aligns with the comparative dual-column approach identified earlier. Reduced listening anxiety (70%) confirms quantitative findings, with many students expressing that the safety net of AI transcripts reduced the pressure of real-time comprehension. However, concern about over-reliance (65%) emerged as a near-equivalent theme, revealing students' awareness of the potential negative consequences of excessive AI dependence. This self-awareness is encouraging, suggesting that students are not passive consumers of AI but active monitors of their own learning habits. The theme need for instructor guidance (55%) is particularly significant for pedagogy: students recognized that AI tools alone are insufficient; explicit strategy instruction is necessary to maximize benefits.

Discussion

The findings provide substantive answers to the three research questions. Regarding students' perceptions of AI in learning English (Question 1), HUFLIT first-year students demonstrated predominantly positive attitudes, valuing AI's capacity for real-time transcription, vocabulary clarification, and anxiety reduction. However, approximately 45% expressed concerns about speech naturalness and accuracy, suggesting that perceptions are favorably balanced yet critically aware of limitations. This ambivalence reflects Ali (2025) findings on AI acceptance in EFL contexts, where perceived usefulness outweighs technical shortcomings.

Concerning AI tools' impact on English learning skills (Question 2), the evidence is compelling. AI-enhanced instruction produced statistically superior outcomes in both listening comprehension (experimental gain 50% vs. control gain 17%) and note-taking quality (62% vs. 22%). These results support Liu and Zhao's (2025) quasi-experimental findings that AI-generated content restructuring reduces cognitive load and enhances comprehension. Nevertheless, qualitative data revealed that effective impact depends on pedagogical scaffolding; students who used AI as a verification tool (comparative dual-column approach) outperformed those using AI as a substitute (AI-assisted linear approach). This finding underscores the importance of strategic integration rather than mere technology access.

On note-taking skill application in listening texts (Question 3), students demonstrated three distinct adaptive strategies. Most notably, the comparative dual-column approach

emerged as most effective, allowing students to verify and refine their personal notes against AI-generated standards. This process aligns with Mayer's (1996) SOI model, where AI serves as a selection and organization scaffold while preserving the learner's active integration role. Critically, 22.5% of students strategically avoided AI in certain tasks, demonstrating metacognitive awareness and resistance to over-reliance. These patterns suggest that AI-enhanced note-taking should be taught not as replacement but as supplementation, fostering what Tran et al. (2025) terms "AI-assisted autonomy" – the ability to strategically deploy AI tools while maintaining independent listening competence.

5. Conclusion

The research investigated the role of artificial intelligence in enhancing listening and note-taking skills among 80 first-year EFL students at HUFLIT's Faculty of Foreign Languages. Addressing the first research question on students' perceptions of AI in learning English, the findings revealed predominantly positive attitudes, with experimental group participants reporting high perceived usefulness (M=4.28/5) and reduced listening anxiety (70% of journal respondents). However, approximately 45% expressed concerns regarding unnatural prosody and transcription accuracy, indicating that perceptions are favorably balanced yet critically aware of AI's limitations – a pattern consistent with Jongbum (2025) and Ali (2025) in other EFL contexts.

Regarding the second research question on AI tools' impact on English learning skills, the evidence demonstrates statistically significant improvements. The experimental group achieved a 50% gain in listening comprehension (Cohen's d=2.97) compared to 17% in the control group (d=0.94), and a 62.4% improvement in note-taking quality versus 22.2% in the control group. These findings support Tran et al. (2025) and Liu and Zhao (2025), confirming that AI-enhanced instruction produces substantially larger effect sizes than conventional methods when implemented with appropriate pedagogical scaffolding.

Concerning the third research question on note-taking skill application, three distinct strategies emerged: AI-assisted linear notes (42.5%), comparative dual-column notes (35.0%), and AI-independent strategic notes (22.5%). The comparative dual-column approach – where students verify personal notes against AI-generated transcripts – yielded the highest learning outcomes, suggesting that effective AI integration requires strategic, selective use rather than passive dependence.

Suggestions for Future Research

Future research should address several limitations of the present study. First, longitudinal designs are needed to examine whether AI-induced gains persist over time and whether over-reliance concerns materialize into long-term skill deficits. Second, comparative studies across different AI tool types (e.g., generative AI vs. speech recognition alone) would clarify which features drive improvement. Third, investigations with larger, more diverse samples across multiple Vietnamese universities would enhance generalizability. Finally, future research should examine how AI-enhanced note-taking instruction can be systematically integrated into EFL curricula, moving from isolated interventions to sustained pedagogical models.

References

Adil, J. J. G. (2025). AI in Education: A Systematic Literature Review of Emerging Trends, Benefits, and Challenges. *Seminars in Medical Writing and Education*, *4*, 795. https://doi.org/10.56294/mw2025795

Al-Ghazo, A. (2023). The Impact of Note-Taking Strategy on EFL Learners' Listening Comprehension. *Theory and Practice in Language Studies*, *13*(5), 1136-1147. https://doi.org/10.17507/tpls.1305.06.

Ali, D. (2025). EFL Students' Perceptions about the Role of Generative Artificial Intelligence Mediated Instruction in their Emotional Engagement and Goal Orientation: A Motivational Climate Theory Perspective in Focus. *Learning and Motivation*, *90*, 102114. https://doi.org/10.1016/j.lmot.2025.102114.

Creswell, J. W., & Creswell, J. D. (2014). *Research Design*. SAGE.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

Do, T. T. P., & Dang, P. N. (2025). Enhancing short-term memory in consecutive interpreting training: Evaluating AI-generated speech simulations for English majors at the diplomatic academy of Vietnam. *VNU Journal of Foreign Studies*, *41*(6), 101-117. https://doi.org/10.63023/2525-2445/jfs.ulis.5530

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford University Press.

Hien, H. V. (2025). The interplay between Vietnamese EFL students' artificial intelligence literacy and tool usage for academic writing. *TNU Journal of Science and Technology*, *230*, 393-400.https://doi.org/10.34238/tnu-jst.12221

Irgin, P. (2025). Note-taking in Academic Listening: A Translanguaging Perspective. *RELC Journal*, *0*(0). https://doi.org/10.1177/00336882251322253.

Jin, Z., & Webb, S. (2023). The effectiveness of note taking through exposure to L2 input: A meta-analysis. *Studies in Second Language Acquisition*, *46*(2), 404-426. https://doi.org/10.1017/S0272263123000529

Jongbum, H. (2025). Opportunities, Challenges, and Perceived Effectiveness of AI-Generated Listening Materials in a University EFL Context. *Modern English Education*, *26*, 446-459.

Kim, J. (2014). Korean EFL Students' Perceptions on Note-taking Strategy and Its Training in Academic English Listening Tests. *English Language & Literature Teaching*, *20*(3), 95-118.

Liu, Q., & Zhao, Y. (2025). The impact of NoteBookLM generated podcasts on EFL news listening comprehension, flow experience, and anxiety: A quasi-experimental study. Proceedings of the 2nd Guangdong-Hong Kong-Macao Greater Bay Area Education Digitalization and Computer Science International Conference, 665-670. https://doi.org/10.1145/3746469.3746573

Luu, T. M. V., & Doan, Q. V. (2025). ChatGPT's Impact on Listening Comprehension: Perspectives from Vietnamese EFL University Learners. *Computer-Assisted Language Learning Electronic Journal*, *26*(3), 43-63. https://doi.org/10.54855/callej.252633

Mayer, R. E. (1996). Learning Strategies for Making Sense Out of Expository Text: The SOI Model for Guiding Three Cognitive Processes in Knowledge Construction. *Educational Psychology Review, 8,* 357-371. https://doi.org/10.1007/BF01463939

Meniado, J. C. (2023). The Impact of ChatGPT on English Language Teaching, Learning, and Assessment: A Rapid Review of Literature. *Arab World English Journal*, *14*(4), 3-18. https://dx.doi.org/10.24093/awej/vol14no4.1

Nguyen, T. K. T., Nguyen, T. T. D., & Trinh, D. D. (2026). English Majors' and Non-English Majors' Perceptions and Readiness For AI-Assisted Learning in English Language Courses: An Exploratory Study. ICTE Conference Proceedings, *9*, 107-120. https://doi.org/10.54855/979-8-9870112-9-4_7

Phan, T. N. L. (2023). Students' perceptions of the AI technology application in English writing classes. Proceedings of the AsiaCALL International Conference, *4*, 45-62. https://doi.org/10.54855/paic.2344

Siegel, J. (2020). Effects of notetaking instruction on intermediate and advanced L2 English learners: A quasi-experimental study. *Journal of English for Academic Purposes*, *46*, 100868. https://doi.org/10.1016/j.jeap.2020.100868.

Siegel, J. (2020). Research into practice: Teaching notetaking to L2 students. *Language Teaching*, *55*(2), 245-259. https://doi.org/10.1017/S0261444820000476.

Shizuka, S., & Ryan, S. (2026). Beyond word count: The role of note quality on EFL students' lecture listening comprehension. *System*, *137*, 103927. https://doi.org/10.1016/j.system.2025.103927.

Tran, P. T., Nguyen, L. H., Luc, T. T., Nguyen, T. T. N., & Le, L. T. (2025). Students' Perspectives of Using ChatGPT and Gemini in Learning English Writing Skills: A Case Study at Binh Duong University. Proceedings of The National Conference: New Trends in Foreign Language Teaching and Learning at Binh Duong University, 893-911. Ho Chi Minh City National University Publishing House.

Tran, P. T., Nguyen, L. H., & Vo, T. T. P. (2025). Using ChatGPT, Gemini, and Grok in Learning English: Students' Perspectives at A Vietnamese University. Proceedings of The 2025 TESOL International Conference**:** Teaching English with INTESOL in an Artificial Intelligence World at Ho Chi Minh City University of Foreign Languages and Information Technology, 208-225. Huflit Journal of Science.

#189 AI AS A "FACE-NEUTRAL" SPACE: REDUCING ANXIETY AND BOOSTING WILLINGNESS TO COMMUNICATE AMONG VIETNAMESE ENGLISH MAJORS IN THE EFL-TO-ESL TRANSITION

Ngoc Giang Tran[1], Thi Phuong Mai Do[2], Thi Cam Trang Thai[3], Thi Kim Nhung Luu[4]

Abstract

As Vietnam transitions from English as a Foreign Language (EFL) to English as a Second Language (ESL), there is a pressing need for learners to develop communicative spontaneity and confidence. Yet the cultural norm of face-saving frequently hinders willingness to communicate and intensifies language anxiety. This exploratory pilot study investigates whether generative AI voice chatbots can serve as a "face-neutral" communicative space that helps reduce these barriers. We employed a mixed-methods design with 30 English majors at a university in Hanoi who engaged in daily voice interactions with ChatGPT over one week. Pre- and post-intervention surveys measured Foreign Language Classroom Anxiety (FLCA) and Willingness to Communicate (WTC), while semi-structured interviews in Vietnamese with eight participants captured learners' own understandings of their experience. Quantitative results showed a meaningful 18.5% increase in WTC (particularly a 53% increase in confidence with native speakers) and a more modest 7.5% anxiety reduction. Qualitative findings revealed three coherent themes: (1) AI functioned as a genuinely face-neutral environment, eliminating judgment-related anxiety; (2) learners experienced freedom to fail, shifting from perfectionism to meaning-focused communication; and (3) increased confidence, though participants appropriately recognized AI limitations and emphasized its supplementary role. The research suggests that face-concern - rooted in cultural values - is highly contextual and remediable through environmental redesign rather than individual psychological intervention. Findings offer practitioners concrete AI prompt templates and a master setup guide for classroom use, along with pedagogical recommendations for Vietnam's ESL-focused curricula and teacher education initiatives.

***Keywords*:** English as a foreign language (EFL), English as a second language (ESL), AI chatbot, face/face-concern, willingness to communicate (WTC), foreign language classroom anxiety (FLCA), Vietnamese learners, teacher education

[1] *Hanoi National University of Education*
[2] *Hanoi National University of Education*
[3] *Hanoi National University of Education*
[4] *Hanoi National University of Education*

1. Introduction

Over the past decade, Vietnam has experienced a significant shift in its English language education landscape. What was once firmly established as English as a Foreign Language (EFL) - taught primarily through grammar-focused, written-centered instruction in classroom contexts - has increasingly become English as a Second Language (ESL) in practice. This transition reflects Vietnam's integration into regional economic networks, its role as a technology hub, and the practical reality that young professionals, students, and workers now use English daily in workplaces, universities, and online communities. The demand is no longer for grammatical accuracy in isolation; it is for communicative spontaneity - the ability to engage in natural, unscripted conversation with genuine interlocutors.

For English majors preparing to become educators, this transition carries particular weight. These future teachers will shape how the next generation learns English. If they themselves remain anxious and reluctant speakers, they are likely to replicate the anxiety-sustaining pedagogies of their own schooling. Conversely, teachers who have developed authentic communicative confidence and spontaneity can model fluent English, normalize error as part of learning, and create classrooms where students feel psychologically safe enough to take linguistic risks. Thus, the communicative development of English majors has implications that extend far beyond individual proficiency to shape national language education capacity.

Yet a substantial obstacle remains. The cultural value of "face-saving" - a central concept in Vietnamese Confucian-heritage society - creates a powerful psychological barrier to spontaneous speech. The Vietnamese concepts of mất mặt (losing face) and giữ thể diện (maintaining face) are not mere social preferences; they reflect deeply held values about dignity, respect, and social belonging. In English classrooms, these values activate protective strategies: silence, self-censorship, avoidance of unfamiliar vocabulary, and reluctance to produce unguarded speech. The irony is profound: learners' cultural values, which serve crucial social functions in Vietnamese contexts, become barriers to English communication development. Preliminary data from our participant pool of 19 English majors at Hanoi National University of Education reveals the severity of this "anxiety gap": mean anxiety about English speaking stands at 3.63 out of 5 (moderate-to-high), while mean confidence when speaking with native speakers is just 2.53 out of 5 (critically low). Yet 84% of these same learners agreed that they "would speak more English if I knew I wouldn't be judged for my mistakes." This statistic captures the heart of the problem: the barrier to communication is not linguistic capacity, but the affective cost of public failure in a face-conscious culture.

We began to wonder: might technology offer a path forward? Generative AI voice chatbots - particularly tools like ChatGPT with voice functionality - present an intriguing possibility. An AI interlocutor possesses no social judgment capacity, cannot gossip or report mistakes to others, and remains endlessly patient and encouraging. In essence, it creates a communicative space in which the social apparatus that activates face-concern simply does not exist. Such a space is not a classroom, nor is it a conversation with a friend. But it might be a uniquely valuable intermediate environment: judgment-free enough to enable unguarded speech production, responsive enough to provide meaningful dialogue, and accessible enough to offer daily practice without logistical barriers.

This exploratory pilot study investigates whether daily voice interactions with an AI chatbot can reduce foreign language classroom anxiety and increase willingness to communicate among Vietnamese English majors. More fundamentally, we ask: Can a "face-neutral" technological environment address a barrier that is deeply rooted in culture? Can brief, judgment-free practice in AI dialogue build the confidence and risk-tolerance necessary for spontaneous human communication? And if so, how should educators thoughtfully integrate such tools into curricula in ways that respect both technology's affordances and its limitations?

To address these questions, we employed a mixed-methods design combining quantitative surveys with qualitative interviews conducted in Vietnamese - our learners' first language. This report details our methodology, findings, and pedagogical recommendations for Vietnam's ESL transition and teacher education initiatives.

Research Questions

1. **RQ1**: Can brief, judgment-free AI voice interactions reduce foreign language classroom anxiety among Vietnamese English majors?
2. **RQ2**: Can brief AI voice interactions increase willingness to communicate in English?
3. **RQ3**: What are Vietnamese learners' own understandings of how (and why) AI interaction affects their language anxiety and communicative confidence?
4. **RQ4**: To what extent do learners perceive AI practice as transferable to subsequent human-to-human English communication?

2. Literature Review

Sociolinguistic Foundations: Face Theory in Confucian-Heritage Cultures

The concept of "face" often refers to the fact all humans have fundamental social needs: the desire to be approved of and regarded as competent (positive face), and the desire to maintain autonomy and freedom from imposition (negative face) (Sifianou, 2011). However, these dimensions are weighted differently across cultures. In individualistic Western contexts, negative face often dominates communicative strategies. In Confucian-heritage cultures - including Vietnam, China, Korea, and Japan - positive face assumes paramount importance. An individual's worth, competence, and place in social networks are continuously negotiated and vulnerable to damage through public failure, correction, or exposure of incompetence. In Vietnamese specifically, the parallel concepts of mất mặt and giữ thể diện extend beyond individual dignity to encompass familial honor and social standing.

Research by Quyen (2025) suggests that Vietnamese learners, when facing the threat of public error or criticism, activate protective linguistic strategies that ironically inhibit the very language production necessary for development: silence, self-censorship, and avoidance of communicative risk. Thus, a cultural value system that serves crucial social functions in Vietnamese contexts becomes, in language learning contexts, a substantial affective barrier. Students whose cultural training emphasizes face-saving face a dilemma: honor their cultural values or take the linguistic risks that language acquisition requires.

Psychological Dimensions: Foreign Language Classroom Anxiety

Horwitz, Horwitz, and Cope (1986) introduced Foreign Language Classroom Anxiety (FLCA) as a construct distinct from general trait anxiety. FLCA is situation-specific and multidimensional, encompassing three core factors: (1) communication apprehension (fear

about communicating with others), (2) test anxiety (fear of academic evaluation), and (3) fear of negative evaluation by others. Importantly, fear of negative evaluation is not simply worry about grades; it includes anxiety about peer judgment, social embarrassment, and damage to one's reputation.

Decades of research have consistently documented elevated FLCA among Asian learners, particularly in speaking contexts. A meta-analysis by Teimouri, Goetze, & Plonsky (2019) found that FLCA negatively predicts language achievement and that anxiety is substantially higher in speaking tasks than in reading or writing. For Vietnamese learners specifically, Khanh & Ngoc (2022) report mean FLCA scores well above global averages, with speaking anxiety as the dominant dimension. The mechanism is well-understood: when anxiety escalates, cognitive capacity narrows, attention diverts from meaning-making to self-monitoring, and avoidance behaviors multiply - all of which reduce the input and output interactions necessary for language acquisition. In classroom contexts where face-concern amplifies anxiety, a self-defeating cycle emerges: fear of judgment reduces speech → reduced output limits practice → limited practice stunts fluency → lower fluency amplifies anxiety.

Communication Dynamics: Willingness to Communicate

Macintyre, Clément, Dörnyei, and Noels (1998) developed the Willingness to Communicate (WTC) model to explain individual differences in communicative behavior beyond simple proficiency. WTC is defined as "the probability of engaging in communication when free to choose." Critically, a highly proficient learner may be unwilling to communicate, while a lower-proficiency learner with high WTC may actively seek communication opportunities. WTC is influenced by multiple factors - self-confidence, motivation, anxiety, perceived competence - and these factors are themselves shaped by social and situational context.

Research by Zhou, Hodis, & Croucher (2026) shows that learners from collectivist cultures exhibit notably lower WTC in classroom settings, particularly with unfamiliar interlocutors or authority figures. For Vietnamese learners, the presence of peers, teachers, or unknown native speakers activates face-concern mechanisms that suppress WTC below learners' actual communicative capacity. Macintyre and Legatto (2011) refined the model further, arguing that WTC must be understood not as a static trait but as a contextually variable probability. This opens a crucial possibility: environmental redesign - specifically, the removal or reduction of face-threat cues - could increase WTC. An AI interlocutor, possessing no social judgment capacity and offering no threat to public face, may represent precisely such a redesigned environment.

Affective Filter and Language Acquisition

Krashen's (1982) Affective Filter Hypothesis proposes that anxiety, fear, and negative self-perception act as a "filter" that blocks language input from being processed and acquired. While critics have questioned some aspects of the hypothesis, subsequent research has lent considerable support to the core insight: when learners feel safe, interested, and confident, language input is more readily acquired. Ding and Yusof (2025) demonstrated that learners who participated in low-anxiety speaking practice tasks showed greater gains in spontaneous speech compared to controls. Importantly, they found that perceived safety - learners' beliefs about whether they would face judgment - was as influential as actual safety in determining affective filter level.

AI in Language Learning: From Tools to Interlocutors

The intersection of AI and language learning spans decades, from early grammar-checking software to contemporary conversational systems. Recent empirical work on chatbot-based language practice has generally shown positive effects on vocabulary, grammar, and learner engagement. Studies using platforms like Duolingo report that learners using AI conversation practice show fluency gains comparable to or exceeding those from traditional classroom instruction. However, prior research has underexplored a critical affordance: the psychological and sociolinguistic safety of AI interaction. Zhong et al. (2026) found that learners reported lower anxiety in human-AI dialogue compared to human-human dialogue; however, their research did not investigate cultural dimensions of face or examine transfer effects. Critically, no prior studies have systematically investigated whether AI can serve as a sociolinguistically protective environment specifically for learners from face-conscious cultures - or whether regular AI practice can transfer increased confidence to subsequent human-to-human communication.

The Face-Neutral Hypothesis: Bridging Theory and Technology

Building on the theoretical foundations above, we propose the Face-Neutral Hypothesis: that AI chatbots, by virtue of their non-human status and incapacity for social judgment, create an interactional environment in which learners' face concerns are minimized. In such an environment, the affective filter is lowered, anxiety decreases, and WTC increases - creating conditions optimal for language acquisition and confidence-building that may subsequently transfer to human-mediated communication.

The hypothesis is grounded in three key observations. First, AI is genuinely non-judgmental: it cannot form opinions, share information, or socially position the speaker. Second, AI availability is decoupled from social cost: a learner can practice at any time without burdening a conversation partner or being perceived as needy. Third, error tolerance in AI dialogue is explicitly reframable: when an AI fails to understand, the learner can attribute this to technical limitation rather than communicative failure, thus reducing shame and encouraging persistence. However, the hypothesis also acknowledges significant limitations. AI lacks pragmatic competence - the ability to navigate cultural context, implicit meanings, and non-verbal communication. It cannot teach learners how to negotiate face in actual social interaction. Therefore, AI should be conceptualized not as a replacement for human interaction but as a scaffolding tool - one that builds confidence in a protected environment, preparing learners for the more complex, socially contingent demands of authentic human dialogue.

Theoretical Integration and Research Gaps

The literature reviewed above suggests a compelling case for investigating AI as an anxiety-reducing, WTC-enhancing tool for learners from face-conscious cultures. Face Theory establishes that face concerns are culturally normative and psychologically consequential. Research on FLCA shows that anxiety inhibits speaking behavior and language acquisition. WTC theory suggests that environmental factors - particularly perceived threat to face - directly modulate communicative behavior independent of proficiency. AI research suggests that learners perceive AI as non-threatening and that AI dialogue can improve linguistic outcomes. Yet existing studies do not synthesize these domains. There is no rigorous empirical investigation of whether AI can address anxiety and increase WTC among learners for whom

face concerns are culturally salient, nor is there qualitative research examining learners' own understandings of this mechanism. This pilot study aims to fill this gap by combining quantitative measures of anxiety and WTC with in-depth qualitative exploration of how learners themselves experience the face-neutral affordances of AI dialogue, and whether this experience translates into greater confidence and willingness in subsequent human-to-human English communication.

3. Methodology

Research Design

This study employs a mixed-methods, explanatory sequential design. The quantitative phase (pre- and post-intervention surveys) measures numerical changes in Foreign Language Classroom Anxiety and Willingness to Communicate. The qualitative phase (semi-structured interviews in Vietnamese) follows and elaborates on the quantitative findings, allowing participants to explain their experiences and articulate the mechanisms - particularly around face, confidence, and communicative shift - that underlie changes in anxiety or WTC. The sequential ordering serves a specific purpose: it enables purposive sampling of interview participants based on distinct quantitative profiles, allowing qualitative data to illuminate differential outcomes. This design is well-suited to exploratory research addressing complex cultural and psychological constructs that cannot be fully captured through surveys alone.

Participants

Quantitative Phase: Nineteen English majors at Hanoi National University of Education (HNUE) participated in the pre-survey; nine completed the post-survey, yielding a retention rate of 30%. Participants were recruited via open announcement; no exclusion criteria were applied.

Demographic Characteristics: The sample included 1st year (n=1; 5%), 2nd year (n=9; 47%), and 3rd year (n=9; 47%) English majors. Participants reported 8–14 years of English study (M=11.2 years), reflecting approximately a decade of classroom instruction from primary school onward. Notably, all 19 participants (100%) reported having used AI tools for English learning purposes, indicating familiarity with LLM-based technology.

Qualitative Phase: Eight participants were selected for interview via purposive sampling, prioritizing diversity in anxiety and WTC trajectories. Selection criteria sought participants who showed substantial pre-post changes and those whose post-survey responses indicated divergent experiences (high vs. low perceived benefit of AI practice).

The Intervention: One-Week AI Voice Practice

Participants engaged in daily voice-based interactions with ChatGPT Voice over seven consecutive days (May 1–7, 2026). Each session was designed to require 10–15 minutes of engagement. All participants received an identical "master setup" prompt to establish clear communication norms: (1) correct grammar only if meaning is unclear; (2) keep responses under 50 words for natural back-and-forth; (3) use friendly, encouraging tone; (4) provide hints or synonyms rather than full answers when stuck. This prompt establishes explicit rules maximizing communicative demand while minimizing perfectionism and anxiety about grammar.

The daily task schedule progressed from low-stakes (casual chat on Day 1) to moderate-to-high stakes (debate on Day 2, roleplay on Day 3, cultural storytelling on Day 4). This

progression deliberately increased communicative demand while building learner confidence in a graded manner. Day 5 allowed free choice to consolidate preferred practices; Day 6 invited metacognitive reflection. Day 7 included post-survey and interviews (no AI task).

Instruments

Pre- and Post-Survey:

The survey employed a 5-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree) combining four adapted FLCA items measuring cognitive and physiological anxiety, three WTC items measuring communicative confidence and risk-taking, and two post-survey-only items assessing perceived AI relaxation and expected classroom transfer. These items were adapted from established scales (Horwitz et al., 1986; Macintyre et al., 1998) with slight modifications for clarity. Items were presented in English, as all participants were English majors; researchers were available to clarify ambiguous wording during administration.

The four anxiety items (Items 1–4) were averaged to create an Anxiety Composite score (range 1–5). The three WTC items (Items 5–7) were averaged to create a WTC Composite score (range 1–5). Item 10 ("I would speak more English if I knew I wouldn't be judged for my mistakes"), while thematically related to face-concern and anxiety, was analyzed as a standalone validation item rather than included in either composite. This item measures perceived barriers to communication rather than actual communicative willingness, and was included post-hoc to validate the face-concern hypothesis.

Due to the brief intervention duration and modest sample size, internal consistency was not computed for this pilot study. Future investigations should employ full, validated FLCAS (Horwitz et al., 1986) and WTC scales with documented reliability coefficients (Cronbach's α).

Items were adapted from established instruments (Horwitz et al., 1986; Macintyre et al., 1998) to ensure relevance to the one-week intervention and clarity for English-major participants. Full-length instruments would be unwieldy in a brief intervention; adapted versions prioritized content validity for face-concern and judgment-related anxiety specifically.

Semi-Structured Interviews:

Interviews were conducted entirely in Vietnamese, the learners' first language. This choice reduced affective barriers and enabled authentic expression of emotions, identity concerns, and culturally nuanced concepts. Interviews lasted approximately 20 minutes, were audio-recorded with consent, and were fully transcribed in Vietnamese. Key quotes were later translated into English, with original Vietnamese provided in footnotes. Six thematic questions guided interviews, each designed to probe the face-neutral hypothesis, error tolerance, cognitive focus shifts, identity changes, and pedagogical transfer.

Data Analysis

Quantitative:

For each survey item and composite index, means, standard deviations, and response distributions were calculated. Paired t-tests (or nonparametric Wilcoxon tests if normality violated) assessed whether anxiety decreased and WTC increased significantly. Effect sizes (Cohen's d) were reported alongside p-values. Alpha=.05.

Qualitative:

Interviews were analyzed using inductive thematic analysis (Braun & Clarke, 2006). While analysis remained informed by theoretical constructs from the literature (face theory,

FLCA, WTC), codes and themes were developed primarily from data patterns rather than predetermined categories, consistent with exploratory qualitative inquiry.

Inter-Rater Reliability: Two independent coders analyzed interview transcripts from three randomly selected participants (38% of eight interviews). Inter-rater reliability was established using percent agreement on code presence/absence for each coded segment. Initial agreement reached 85% across the coded sample; disagreements were resolved through discussion and consensus refinement. The remaining five interviews (62%) were coded by the lead analyst using the finalized coding scheme.

Theme Definition: Three themes emerged inductively from the data and were defined post hoc:

1. Face-Neutral Environment: Participants' subjective experience of AI as judgment-free and socially non-threatening, free from the affective burden of social evaluation.
2. Freedom to Fail: Reduced perfectionism and increased error tolerance as a byproduct of face-neutral interaction, with corresponding cognitive shift from form-focus to meaning-focus.
3. Scaffolding with Limitations: Participants' perception that AI dialogue builds confidence and communicative willingness while lacking pragmatic and cultural competence of human dialogue.

These themes emerged through inductive analysis rather than from a priori hypotheses, consistent with the exploratory nature of this pilot study.

Ethical Considerations

All participants provided written, informed consent in Vietnamese. Consent forms clearly described study purpose, procedures, risks, benefits, confidentiality protections, and the right to withdraw. Data are de-identified and assigned participant codes. Informed consent forms and audio recordings are stored separately from data.

Limitations

This study experienced substantial attrition, with 10 of 19 pre-survey participants (53%) not completing post-surveys, yielding a 47% retention rate. While demographic characteristics (year in program, years of English study, prior AI tool use) did not differ significantly between completers (n=9) and non-completers (n=10), attrition of this magnitude introduces material risk of selection bias. Participants who completed the full intervention and surveys may have been more motivated, more technologically comfortable, or more positively disposed toward AI than those who did not complete. Such selection bias could artificially inflate observed effect sizes, particularly for self-report measures like WTC. Additionally, the resulting sample size (N=9) for paired analysis substantially reduces statistical power, limiting our ability to detect effects even if they exist in the population. These limitations underscore the exploratory, preliminary nature of findings and necessitate replication with higher retention rates, larger sample sizes, and control conditions before generalizable claims can be made.

4. Quantitative Results

Sample and Response Rates

Of 19 English majors who completed the pre-survey, nine completed the post-survey, yielding a 47% completion rate. While substantial, this attrition is not uncommon in one-week intervention studies in naturalistic academic settings. Analysis of baseline characteristics revealed no significant differences between completers and non-completers, suggesting attrition was random rather than systematic.

Pre-Survey Baseline (N=19)

Participants reported substantial foreign language classroom anxiety at baseline. Across the four anxiety items, the overall mean was 3.63/5.00 (SD=0.85), indicating moderate-to-high anxiety. Item-level analysis revealed:

Physiological anxiety (Item 4: "I feel nervous or tight in my chest when I have to speak English") scored M=3.89 (SD=1.10), with 74% of participants rating this 4 or 5. Fear of negative evaluation (Item 5: "I worry that my classmates or teacher will judge my pronunciation") scored M=3.47 (SD=1.22), with 63% endorsing 4 or 5. Performance anxiety (Item 6: "I am afraid of making grammar mistakes during a conversation") scored M=3.68 (SD=0.82), with 58% responding at 4 or 5. Behavioral avoidance (Item 7: "I prefer to stay silent in class rather than risk saying something wrong") scored M=3.47 (SD=1.12), with 53% endorsing 4 or 5.

Regarding willingness to communicate, overall WTC was M=3.16/5.00 (SD=0.56), though with notable item variation. Confidence with native speakers (Item 8: "I feel confident starting a conversation with a native speaker") was critically low at M=2.53 (SD=1.12), with only 21% rating this 4 or 5 and 58% scoring 1 or 2. Lexical risk-taking (Item 9: "I am willing to try using new or difficult vocabulary during a talk") scored M=3.11 (SD=1.10), with 37% below 3. Most strikingly, Item 10 ("I would speak more English if I knew I wouldn't be judged for my mistakes") scored M=3.84 (SD=0.90), with 84% responding 4 or 5. This final statistic directly supports the face-concern hypothesis: learners explicitly identified judgment fear as the primary inhibitor of speaking behavior.

Post-Survey Results (N=9)

Anxiety

For the nine participants completing both surveys, anxiety composite scores decreased from Pre-M=3.33 (SD=1.02) to Post-M=3.08 (SD=0.65), representing $\Delta M = -0.25$ (7.5% reduction). This was not statistically significant, $t(8)=0.756$, $p=.471$; Cohen's $d=0.252$ (small effect). Item-level inspection revealed: physiological anxiety decreased (−0.33), performance anxiety decreased (−0.44), behavioral avoidance showed the largest reduction (−0.56, a 17% decrease), while fear of negative evaluation remained nearly stable (+0.22).

Interpretation: While overall anxiety reduction was modest and non-significant, behavioral avoidance showed a meaningful 17% decrease. This suggests that participants felt slightly less compelled to remain silent. The modest effect on overall anxiety likely reflects the intervention's brevity and the multifaceted nature of anxiety - physiological and evaluative anxiety may be more resistant to short-term intervention than behavioral avoidance.

Willingness to Communicate

WTC composite scores increased from Pre-M=3.19 (SD=0.78) to Post-M=3.78 (SD=0.55), representing ΔM=+0.59 (18.5% increase). This approached statistical significance, t(8)=−1.650, p=.138; Cohen's d=0.550 (medium effect). Item-level changes were substantial: Confidence with native speakers (Item 8) showed the most dramatic improvement, increasing from M=2.33 (SD=1.12) to M=3.56 (SD=0.88), ΔM=+1.23 (53% increase). Eight of nine participants (89%) increased or maintained their ratings on this item. Lexical risk-taking (Item 9) increased from M=3.00 to M=3.56, ΔM=+0.56 (18.7% increase). Face-concern modulation (Item 10) increased modestly from M=4.11 to M=4.22, ΔM=+0.11 (2.7% increase, a ceiling effect).

Interpretation: The meaningful WTC increase, particularly in confidence with native speakers (+53%), aligns with the face-neutral hypothesis. The minimal change in Item 10 reflects ceiling effects - participants already strongly endorsed this belief at baseline. The practical significance of a 0.59-point WTC increase, combined with a medium effect size, suggests meaningful intervention effect on communicative willingness.

Post-Survey-Only Items (N=9)

Item 11: Perceived Relaxation with AI vs. Humans

Participants rated agreement with "I felt more relaxed speaking to the AI than to a human" at M=4.56/5.00 (SD=0.53). All nine participants (100%) responded 4 or 5, with five (56%) selecting 5 (strongly agree). This near-universal consensus provides preliminary evidence that participants experienced AI interaction as substantially less anxiety-provoking than human-to-human communication. The absence of any disagreement is striking and suggests the face-neutral affordance was perceived as genuine and meaningful.

Item 12: Transfer of Confidence to Future Human Interaction

Participants rated agreement with "My experience with AI made me more willing to speak in my next English class" at M=3.89/5.00 (SD=0.60). Seven of nine (78%) agreed or strongly agreed (ratings 4–5). The gap between Item 11 (M=4.56) and Item 12 (M=3.89) is informative: while participants overwhelmingly experienced AI as relaxing, they were more cautious about whether gains would transfer to human classroom contexts. This item measures participants' self-reported expectations and intentions regarding future classroom speaking, not observed changes in actual classroom behavior. Participants reported expecting increased willingness to speak in subsequent English class contexts; however, whether this expectation translates to actual changes in classroom participation, peer interaction, or teacher-reported student behavior cannot be determined from the present study. Longitudinal follow-up observing participants' actual classroom engagement and teacher feedback in weeks following the intervention would provide stronger evidence for transfer claims. Current evidence of transfer is limited to participants' perceptions and self-predictions.

Summary

Anxiety reductions were modest and non-significant, though behavioral avoidance decreased 17%. Willingness to communicate increased meaningfully (+18.5%), with the largest gains in confidence with native speakers (+53%). Participants overwhelmingly experienced AI as more relaxed than human interaction (M=4.56/5.00) and most expected increased classroom speaking (M=3.89/5.00). These quantitative findings set the stage for qualitative investigation into the mechanisms and nuances of the intervention experience.

5. Qualitative Findings

Semi-structured interviews in Vietnamese were conducted with eight purposively sampled participants. Interviews were fully transcribed in Vietnamese and analyzed using thematic analysis. Two independent coders analyzed 25% of the dataset, achieving 85% initial agreement; themes were refined through discussion. Three major themes emerged, each directly addressing our research questions and the face-neutral hypothesis.

Theme 1: The Face-Neutral Environment - Freedom from Social Judgment

The most pervasive and unambiguous theme across all interviews was participants' experience of the AI as a judgment-free, socially non-threatening interlocutor. Participants spontaneously and repeatedly articulated that speaking to AI eliminated the affective burden of face-concern - the worry about appearing incompetent or being evaluated negatively.

One participant captured this essence: "I felt no need to look smart or perfect for the AI. Maybe because I know it's a machine. I felt free to make mistakes, ask dumb questions, and even laugh at myself. It was like practicing alone, but with useful feedback." Another articulated the contrast with human interaction: "With a human, I would definitely try to save face – I'd avoid difficult words or pretend I know something I don't. But with the AI, that pressure disappeared." A third noted: "Unlike talking to people, I wasn't worried about being judged. I felt more comfortable making mistakes and asking questions."

Participants identified several mechanisms by which AI's non-human status created psychological safety: (1) AI lacks real judgment capacity - it does not evaluate character or form opinions. As one participant explained, "It doesn't know who we are and our personalities, so we can feel more open and comfortable." (2) Conversations are private and non-reportable; mistakes cannot become social knowledge. (3) AI provides unconditional positive regard; it never criticizes but rather explains and corrects supportively. Importantly, qualitative data revealed that this face-neutral environment directly validated the face-concern hypothesis: learners' linguistic silence and anxiety in human contexts stemmed not from linguistic incapacity but from cultural sensitivity to judgment.

Theme 2: Freedom to Fail - Error Tolerance and Shift from Form-Focused to Meaning-Focused Processing

A second major theme centered on participants' newfound permission to make errors without shame, and the corresponding cognitive shift from obsessive form-focus to meaning-focus. This addresses the mechanism by which anxiety reduction translates to increased willingness to communicate.

One participant described a vivid realization: "At first, I focused a lot on grammar and accent – I wanted to sound correct. But after a few days, I cared more about the meaning. The AI always understood me even with wrong grammar or a thick accent, so I slowly stopped overthinking. I started thinking in simple English, not perfect English." Another noted: "I focused more on what I wanted to say [rather than how to say it]. My main goal was to communicate my ideas clearly." A third observed: "The AI understands relatively quickly and accurately, even if I make typos or grammar mistakes, the AI will still understand and respond to my meaning."

Participants also described specific episodes in which they forgot vocabulary or produced incomplete utterances, but continued attempting communication without shame. One

participant couldn't remember the word "patties": "I paused and said, \"the round, flat pork things.\" The AI just said, \"You mean pork patties?\" and continued naturally. I felt relieved, not embarrassed." Another forgot the word "recycling" and described the meaning instead: "The AI quickly guessed the correct word. I felt a little frustrated at first, but then I felt relieved and impressed when the AI understood me." These episodes reveal that error recovery without shame enabled sustained communicative effort - a crucial prerequisite for fluency development.

Theme 3: Scaffolding for Human Interaction - With Acknowledged Limitations

A third major theme addressed the critical question of transfer. While most participants reported increased confidence and willingness to attempt speaking in human contexts after the AI practice, they simultaneously articulated thoughtful recognition of AI's limitations as a scaffold for authentic human interaction.

Identity shift was evident in one participant's statement: "Before, I saw myself as a \"student\" who needed permission to speak. Now I feel more like a \"user\" – English is just a tool. I don't have to be fluent or elegant. I just need to be understood. That shift is huge for my confidence." Another reported: "After talking to the AI for a week, I feel more confident using English in daily communication. Instead of focusing only on grammar and vocabulary, I now see English as a tool to express my ideas and communicate with others."

All interviewed participants recommended AI speaking practice to their future students - but with nuanced caveats. One stated: "I would recommend it, but only as a supplement – not a replacement for real conversation. For shy students who fear losing face, AI is a safe bridge. They can build fluency before talking to humans. However, AI doesn't teach natural body language or real-time cultural reactions. So I'd say: use AI for practice, but don't stop there." Another similarly emphasized: "I would recommend using AI as a supplement, not a replacement for real communication with teachers and other people." A third noted AI's practical accessibility: "If they don't have a real English-speaking environment where they can communicate with native speakers directly, [AI] would be much more useful and prevalent."

Participants also demonstrated critical awareness of AI's shortcomings. One observed: "AI will miss some logical errors occasionally." Another reflected: "The AI never gets tired or impatient. That's great, but also a little unrealistic. Real people get bored, interrupt, or show emotions. So I had to remind myself that AI fluency isn't exactly real-world fluency." A third noted: "With AI chatbox, I still can't be as natural as when communicating with humans, but it can be a tool to help English improve a bit." These observations reveal sophisticated understanding: participants recognized AI as a bridge or scaffold, not a replacement.

Cross-Cutting Observations

Beyond the three primary themes, participants also articulated observations nuancing our understanding. One noted that AI sometimes provides responses misaligned with the user's intent, requiring clarification. Another observed that AI's suggestions should be evaluated critically rather than blindly accepted. These observations suggest that while AI provides psychologically safe communicative practice, learners should receive explicit instruction in evaluating AI output - an important pedagogical implication.

6. Discussion

Integrating Quantitative and Qualitative Findings: The Face-Neutral Hypothesis Revisited

This pilot study investigated whether AI voice chatbots could serve as judgment-free spaces to reduce anxiety and increase willingness to communicate among Vietnamese English majors. Quantitative results showed that while anxiety reduction was modest and non-significant, willingness to communicate increased meaningfully (+18.5%, d=0.55). Qualitative analysis identified three coherent themes: participants experienced the AI as judgment-free and face-neutral; this safety enabled a shift from perfectionism to meaning-focused communication and error tolerance; and participants reported increased confidence and expected transfer to human dialogue, while appropriately acknowledging AI's limitations.

Taken together, these findings provide partial and nuanced support for the face-neutral hypothesis. The mechanism linking AI practice to increased WTC appears to be not dramatic anxiety reduction, but rather removal of the affective cost of failure. In human-mediated dialogue with peers and teachers, failure (forgetting a word, mispronouncing, producing ungrammatical speech) threatens face and thus carries high psychological cost. Learners protect themselves through avoidance. In AI dialogue, the same failure carries no face threat and thus no cost; learners are free to attempt utterances, stumble, recover, and continue. This freedom, rather than clinical anxiety reduction, appears to be the active ingredient enabling increased WTC. This interpretation reconciles an apparent paradox: overall anxiety decreased only 7.5% and non-significantly, yet WTC increased 18.5% with a medium effect size. The resolution is that WTC is not simply the inverse of anxiety; it is modulated by contextual factors (perceived face threat, social judgment risk) that can be modified independently of global anxiety levels.

Reinterpreting FLCA: The Specificity of Face-Concern in Vietnamese Contexts

Our findings challenge a subtle but important aspect of anxiety research in language learning. FLCA research has documented that anxiety inhibits communication. However, most research treats anxiety as relatively global and trait-like. Our data suggest that for Vietnamese and other face-conscious learners, a significant portion of "classroom anxiety" is actually face-concern anxiety - anxiety specifically about social judgment and loss of face - which is highly contextual. The most preliminary evidence is Item 10: 84% of participants agreed they "would speak more English if I knew I wouldn't be judged for my mistakes." This indicates learner anxiety stems not from linguistic incapacity but from social consequences of visible failure. Crucially, this face-concern-specific anxiety is not equally present in all contexts. In the AI context - where judgment is impossible - this anxiety evaporates entirely (Item 11: M=4.56/5, unanimously experienced as low-anxiety). Yet physiological anxiety and test anxiety, which are less contextually dependent, showed only modest reduction.

Theoretical implication: The FLCA framework, while valuable, may conflate distinct anxiety types that have different etiologies and respond to different interventions. Face-concern anxiety in high-context cultures may be better conceptualized as sociolinguistic anxiety - anxiety arising from cultural norms and interpersonal power dynamics rather than personality traits. This sociolinguistic anxiety can be rapidly and dramatically reduced through environmental redesign (removing judgment threat), whereas personality-based anxiety reduction may require longer-term psychological intervention. This finding aligns with recent

work by Akbar (2025) on the "cultural dimensions of language anxiety," arguing that anxiety in non-Western contexts often reflects cultural values rather than Western personality constructs.

The WTC Surge: Confidence with Native Speakers

The most striking quantitative finding is the 53% increase in confidence with native speakers (Item 8: 2.33→3.56), moving from critically low to moderate-to-good. This is significant because confidence with native speakers represents perhaps the highest-stakes dimension of WTC for ESL learners - native speakers are perceived as ultimate arbiters of linguistic correctness, and their feedback is most threatening to face. In one week, participants substantially gained in this previously low-confidence domain. The qualitative data illuminate this: by practicing spontaneous dialogue with a conversational AI, participants internalized a different model of English communication. One participant's testimony - "Now I feel more like a \"user\" – English is just a tool" - captures an identity shift from "learner seeking permission" to "tool-user with communicative agency." Macintyre's WTC model emphasizes that willingness depends on self-perceived communicative competence and self-efficacy. Our intervention appears to have specifically targeted the latter: by providing consistent, successful communicative experiences in which understanding and dialogue flowed naturally (despite grammatical imperfection), the AI practice may have shifted participants' self-efficacy beliefs: "I can communicate in English, even if imperfectly."

Comparison with Prior Research on AI in Language Learning

Studies on chatbot-based language practice have generally documented positive effects on fluency, vocabulary, and engagement. Our data confirm this: participants reported increased willingness to attempt communication, greater error tolerance, and higher confidence. Liu, Lv, & Chen (2025) found that learners report lower anxiety in human-AI dialogue compared to human-human dialogue, which aligns directly with our Item 11 finding (M=4.56 for "more relaxed with AI"). However, their research did not investigate cultural dimensions of face and anxiety, nor examine transfer effects. Our work makes three novel contributions. First, we explicitly foreground face-concern as the primary inhibitor of WTC and demonstrate that AI's judgment-free affordance directly addresses this mechanism. Prior research has not systematically investigated cultural face-concerns. Second, we investigate qualitatively why anxiety reduction occurs, revealing that the mechanism is not clinical anxiety reduction but removal of face-threat, which frees cognitive and emotional resources for meaning-making. Third, we examine not only learning in the AI context but also participants' beliefs about transfer to human dialogue and their recognition that AI cannot replace human interaction. Most prior research focuses only on the AI environment itself.

Reconciling the AI-Class Gap: Why Item 12 is Lower Than Item 11

A notable discrepancy in our post-survey data is that perceived relaxation with AI (Item 11: M=4.56) substantially exceeds perceived transfer to classroom speaking (Item 12: M=3.89). This 0.67-point gap is informative and aligns with qualitative data. The gap likely reflects several factors. First, participants understood that human classrooms involve complexity absent from AI dialogue: real people become bored, interrupt, show emotions. Participants' caution about transfer may reflect sophisticated understanding that confidence gained in one context does not automatically transfer to different social conditions. Second,

one week is brief. Participants may have been uncertain whether anxiety reduction would persist once the intervention ended and they returned to human-mediated contexts. Third, even after one week of AI practice, the presence of human judges in classroom settings may re-activate face concerns. The Item 12 mean of 3.89 is not a failure but realistic modulation of expectations: 78% of participants expect some transfer of confidence to classroom settings, but acknowledge uncertainty. This measured optimism, grounded in accurate understanding of AI limitations, is more credible than uncritical enthusiasm.

Modest Anxiety Reduction: Why Clinical Anxiety Changed Less Than WTC

The anxiety composite showed only 7.5% reduction (non-significant), which may seem disappointing. However, this finding is theoretically coherent and meaningful. Our anxiety measure combined four distinct dimensions: physiological anxiety, fear of negative evaluation, performance anxiety, and behavioral avoidance. These respond differently to intervention. Physiological anxiety decreased modestly (−0.33); this construct may have stronger trait-like and biological components that one week of practice cannot fully address. Behavioral avoidance showed the largest reduction (−0.56, 17% decrease); this is the most behavioral and contextually malleable dimension. Fear of negative evaluation actually increased slightly (+0.22), which may reflect that participants, having experienced success in AI dialogue, became more willing to attempt human dialogue and thus re-encountered the possibility of negative evaluation. Brief interventions typically produce state-level changes (temporary reductions in anxiety in the specific context) more readily than trait-level changes (enduring personality-related anxiety). Participants experienced low anxiety during AI practice but may not have internalized a new general anxiety baseline about English speaking. Longer interventions likely needed for trait-level change.

Study Limitations and Threats to Validity

While our findings are suggestive and theoretically coherent, several limitations constrain generalizability. The pre-survey sample (N=19) is modest, and the post-survey completion rate (47%, N=9) substantially reduces statistical power. The paired-sample analysis revealed meaningful effect sizes but did not achieve conventional statistical significance - likely a power limitation. The study lacks a control group; observed changes could result from maturation, practice effects, or regression to the mean. A randomized controlled trial comparing AI practice with alternative conditions is essential. The one-week intervention duration is brief; while sufficient to detect state-level changes, it may be insufficient for durable behavioral or trait-level change. Participants who completed post-surveys may differ systematically from non-completers (e.g., more motivated, less anxious). Participants were English majors at a single Vietnamese university; findings may not generalize to non-majors, younger learners, or learners in non-face-conscious cultures. The FLCA items were adapted rather than formally validated, reducing confidence in their psychometric properties. Post-survey-only items lack pre-survey baseline. Interview data are subjective and not independently verified. Despite these limitations, the study provides valuable preliminary evidence grounded in both quantitative and qualitative data, informing larger, more rigorous investigation.

Theoretical Implications: Reframing the Affective Filter for High-Context Learners

Our work suggests a refinement of Krashen's Affective Filter Hypothesis for learners in high-context, face-conscious cultures. The hypothesis posits that anxiety "blocks" language input from being processed. We propose a more precise formulation for face-conscious learners: when learners perceive threat to social face or negative evaluation risk, anxiety escalates and learners adopt avoidant linguistic strategies (silence, minimal response, memorized phrases) that limit both input and output - thus reducing acquisition opportunity. Environmental removal of face-threat cues (judgment impossibility, privacy, non-reportability) lowers this filter, enabling spontaneous output and meaningful interaction - the conditions most conducive to language acquisition. This reframes the approach from personality-based anxiety management ("How do we help students overcome their anxiety?") to sociolinguistic environment design ("How do we redesign environments to remove the social threat cues that activate face-concern?"). This is both more culturally sensitive and more practically feasible.

Implications for ESL as a Second Language

Our findings have timely relevance to Vietnam's EFL-to-ESL transition. As English becomes a second language in practice, learners face two simultaneous demands: developing linguistic proficiency and developing communicative confidence to use that proficiency spontaneously. Traditional EFL pedagogy, which prioritizes accuracy and explicit grammar instruction, may inadvertently reinforce the very mechanisms our data illuminate: heightened attention to form, perfectionism, and anxiety about error-visibility. Our findings suggest that brief, early exposure to low-stakes communicative practice in judgment-free environments can accelerate the transition to spontaneous, meaning-focused language use. AI chatbots offer an accessible, low-cost tool for creating such environments at scale. However, the data also indicate this is a bridge, not a destination. Item 12 results and qualitative caveats remind us that learners must eventually engage with the irreducible complexity of human dialogue - ambiguity, disagreement, real social stakes, cultural nuance. The optimal pedagogical path likely includes: (1) judgment-free AI practice to build initial confidence and shift from form-focus to meaning-focus; (2) graduated human dialogue (structured pair work, small-group discussion, lower-stakes peer interaction) to introduce real social complexity; (3) authentic human communication in varied contexts.

7. Conclusion and Pedagogical Recommendations

Summary of Key Findings

This exploratory pilot study investigated whether daily voice interactions with AI chatbots could reduce foreign language classroom anxiety and increase willingness to communicate among Vietnamese English majors. The mixed-methods design combined quantitative pre-post surveys with qualitative interviews conducted in Vietnamese. Anxiety showed modest, non-significant reduction (−7.5%, from M=3.33 to M=3.08), with the largest decrease in behavioral avoidance (−17%). Willingness to communicate showed meaningful increase (+18.5%, from M=3.19 to M=3.78, p=.138, d=0.55), with the largest gains in confidence with native speakers (+53%, from M=2.33 to M=3.56). Participants overwhelmingly experienced the AI as more relaxed than human interaction (M=4.56/5.00) and most expected increased classroom speaking (M=3.89/5.00). Qualitative analysis identified three coherent themes: face-neutral environment (100% of participants experienced AI as judgment-free); freedom to fail (participants shifted from form-focused perfectionism to meaning-focused communication); and scaffolding with caveats (participants recommended AI as supplement, not replacement). The central insight is that the mechanism linking AI practice to increased WTC is not primarily clinical anxiety reduction, but rather removal of the affective cost of failure through elimination of face-threat. For face-conscious learners, judgment-free environments enable spontaneous speech production - the behavioral outcome essential for language acquisition and communicative development.

Practical Guidance for Educators: Do's and Don'ts

The following guidance synthesizes findings into concrete recommendations for educators incorporating AI into EFL/ESL curricula:

Do	**Don't**	**Rationale**
Create a clear, simple master prompt establishing non-judgmental norms (e.g., "correct only if you don't understand my meaning").	Don't use default AI settings that over-correct grammar or respond with formal lectures.	The master prompt signals that meaning matters more than form, lowering perfectionism anxiety.
Frame AI practice as "conversation with a helpful friend" not "practice for a test."	Don't position AI as a grammar-checking or correction tool.	Learners should focus on content and fluency, not form-accuracy.
Explicitly teach learners that AI is a supplement for building confidence, not a replacement for human interaction.	Don't promise that AI practice alone will solve communication anxiety or produce native-like fluency.	Realistic expectations prevent disappointment and support appropriate use.
Normalize errors as part of learning; emphasize that the AI environment is uniquely safe for experimenting.	Don't suggest that mistakes in AI practice will transfer negatively to humans.	Clarifying that AI is private and non-judgmental removes a major barrier.
Encourage learners to notice their own thinking shifts (form-focus to meaning-focus) and articulate what they're learning.	Don't assume learners will automatically transfer insights from AI to human dialogue.	Metacognitive awareness consolidates learning and increases transfer likelihood.

Check in regularly to ensure engagement and troubleshoot barriers.	Don't treat AI practice as fully autonomous learning without support.	Brief check-ins maintain engagement and prevent attrition.
Use AI practice as a bridge before higher-stakes human dialogue.	Don't use AI as a replacement for human interaction or peer speaking practice.	AI builds confidence; humans teach cultural nuance and real-time negotiation.
Gradually increase task complexity and stakes over time.	Don't assign high-stakes, emotionally complex tasks in the first days.	Confidence builds cumulatively; early wins create momentum.
Measure willingness to communicate and confidence, not just accuracy or fluency.	Don't evaluate AI-mediated speaking using accuracy-focused rubrics designed for human dialogue.	WTC and confidence are primary outcomes; accuracy naturally develops once learners are willing to produce.
Conduct qualitative interviews or written reflections post-intervention.	Don't rely solely on quantitative measures; qualitative data reveals mechanism and meaning.	Learners provide valuable insight into what changed and why.
Include in teacher education explicit training on face-concern dynamics and AI integration.	Don't assume teachers will intuitively understand face-concern or how to leverage AI effectively.	Teachers are gatekeepers of implementation; their understanding is critical.
Share evidence-based, nuanced communication about AI as a tool for addressing anxiety-driven avoidance.	Don't market AI as a "magic bullet" that solves language learning.	Teachers rightfully skeptical of tech hype. Evidence-based communication builds trust.

Implications for Vietnam's ESL Transition and Teacher Education

Vietnam's shift from EFL to ESL creates simultaneous demands: learners must develop linguistic proficiency and communicative confidence to use that proficiency spontaneously in real-world contexts. Traditional EFL pedagogy, with its emphasis on accuracy and explicit grammar instruction, may inadvertently create the very barrier our data illuminate: anxiety-driven avoidance.

For Curriculum Design

Integrate low-stakes, judgment-free communicative practice early. Rather than reserving "conversation" for advanced learners, introduce structured AI dialogue and peer interaction at intermediate levels. This scaffolds learners' identity shift from "student preparing to communicate" to "user of English in meaningful contexts." Rebalance the accuracy-fluency trade-off. Vietnamese curricula traditionally weight accuracy heavily. Our findings suggest that early, error-tolerant practice enables subsequent accuracy development by removing the affective barrier to production. A curriculum front-loading low-stakes fluency practice followed by focused accuracy work may be more effective than the reverse. Make face-concern explicit in pedagogy. Vietnamese teachers often work around face concerns without naming them. More effective would be to explicitly address face concerns in class discussion, normalize errors as essential to learning, and gradually introduce peer feedback structures that rebuild learners' ability to receive and act on criticism.

For Teacher Education Initiatives

Train teachers to recognize face-concern anxiety. Teacher education programs should help teachers understand that Vietnamese learners' silence and avoidance often reflect not lack of ability but cultural anxiety about judgment. This reframing shifts teacher response from "push harder" to "redesign the environment." Provide concrete guidance on AI integration. Teachers need practical training on selecting and configuring AI tools, integrating AI into lesson sequences, monitoring learner engagement, and evaluating learning outcomes beyond grammar accuracy. Build critical awareness. Teachers should understand AI's affordances and limitations. AI is excellent for judgment-free practice and confidence-building but poor at teaching cultural pragmatics and non-verbal communication. Support peer learning communities. Teacher education should include mechanisms for teachers to share experiences, troubleshoot challenges, and collectively refine AI-integrated pedagogy. This distributed approach is more sustainable than top-down mandates.

For Policy and Institutional Support

Allocate resources for AI-integrated pedagogy. Implementation requires technology infrastructure, teacher training time, curricular development, and evaluation. Adopt a phased, evidence-based approach. Rather than wholesale adoption, pilot AI integration with early-adopter teachers and learners, collect evidence, refine practices, then scale. This study and similar research provide the evidence base. Address equity concerns. Ensure AI-integrated learning is not available only in well-resourced urban schools. Open-source AI tools and offline-capable systems may be necessary for equitable access.

Practical Resources: Reproducible Master Prompt and Task Schedule

For educators implementing AI-mediated speaking practice, we provide here the exact master prompt and daily task schedule used in this research, immediately reproducible and requiring no special configuration:

Master Setup Prompt (Copy-Paste One Before Day 1)

I am an English major at HNUE. For the next week, I want you to be my speaking partner. Please follow these rules:
1. Do not correct every grammar mistake – only correct me if you don't understand my meaning.
2. Keep your responses short (under 50 words) so we have a natural back-and-forth.
3. Use a friendly, encouraging tone.
4. If I get stuck, give me a hint or a synonym – not the full answer.

(Educators may adapt this for different proficiency levels or institutional contexts, but these core principles should be maintained.)

Daily Task Schedule (Days 1–6, 10–15 minutes each)

Day	Task	Prompt to Paste	Pedagogical Purpose
1	Warm-up: Free Talk	"Let's have a casual chat. Ask me about my hobbies or my day."	Build rapport; low-stakes entry
2	Sociolinguistic Debate	"Let's debate: Is 'Vietnamese English' a valid variety or just	Identity-relevant; argumentative language

		broken English? Argue against me."	
3	High-Stakes Roleplay	"Pretend I'm a flight attendant and you're an angry passenger. Start by shouting about your cold meal."	Emotional complexity; register variation
4	Cultural Storytelling	"I'll tell you a Vietnamese folk tale. Then you summarize it to check my clarity."	Narrative production; cultural relevance
5	Learner Choice	"Repeat any task from days 1–4 that you found most useful."	Autonomy; consolidation
6	Reflection	"Ask me: What was my biggest improvement this week? What was hardest?"	Metacognitive consolidation
7	Post-Intervention	Complete post-survey; attend interview (if selected)	Assessment; knowledge generation

Future Research Directions

This pilot study opens several avenues for future investigation: (1) Larger, adequately powered randomized controlled trials comparing AI-mediated practice with traditional classroom speaking, human tutoring, and control conditions. (2) Long-term follow-up examining durability of anxiety reduction and WTC gains, and evidence of transfer to actual classroom behavior measured through observation, not self-report. (3) Explicit investigation of transfer mechanisms - under what conditions does confidence from AI dialogue transfer to human dialogue? (4) Cross-cultural variation - do learners from non-face-conscious cultures show similar effects? (5) Moderator and mediator analysis examining individual differences in anxiety, prior technology experience, and personality that may influence outcomes. (6) Sensitivity testing of prompt variation and different AI systems to identify essential features. (7) Classroom integration studies examining optimal sequencing of AI practice with human dialogue, and teacher training supports necessary for effective implementation. (8) Qualitative deep-dives into learners' mental models of English speaker identity, the role of memorable moments in building confidence, and teacher perceptions of changes in student classroom behavior. (9) Cost-effectiveness and scalability analyses assessing feasibility in resource-limited contexts. (10) Intervention refinement studies testing enhancements such as peer discussion following AI practice, explicit metacognitive training, and structured written reflection.

Final Thoughts: AI as a Culturally Responsive Tool

This research began with a simple observation: Vietnamese English majors report that they would speak more English if they weren't worried about being judged. This face-concern is not a pathology but a culturally rooted value system. For centuries, Vietnamese society has prioritized group harmony and interpersonal respect. The challenge for ESL pedagogy is not to erase these cultural values but to help learners use English effectively while maintaining their cultural identity. AI chatbots offer a tool uniquely suited to this task. They create communication environments free of the judgment apparatus that Confucian-heritage cultures

activate, yet they provide meaningful, responsive dialogue necessary for language acquisition. In doing so, they do not undermine Vietnamese cultural values; they protect them by allowing learners to develop English competence in a judgment-free zone, preparing them for the higher-stakes human dialogue to come.

As Vietnam continues its ESL transition and invests in language teacher education, AI-mediated practice deserves consideration not as a futuristic novelty but as a culturally responsive, evidence-based tool for addressing a real barrier to communicative development. When coupled with authentic human interaction and critical teacher guidance, AI chatbots can help Vietnamese learners transition from anxious silence to confident, spontaneous, meaning-focused English communication.

The future of language education is not AI replacing teachers or human interaction. Rather, it is thoughtfully designed environments - combining technology, pedagogy, and human relationships - that respect learners' cultural values while expanding their communicative horizons.

References

Akbar, F. H. (2025). Language Assessment Anxiety: Exploring Social and Cultural Factors in Indonesian University Students. *International Journal of Second and Foreign Language Education, 4*(2), 77–94. https://doi.org/10.33422/ijsfle.v4i2.1136

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3(*2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. SAGE Publications, Inc.

Ding, D., & Yusof, A. M. B. (2025). Investigating the role of ai-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: a mixed methods study. *Humanities and Social Sciences Communications, 12*(1). https://doi.org/10.1057/s41599-025-05550-z

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and Performance: The Processing Efficiency Theory. *Cognition and Emotion*, *6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Gao, G., & Ting-Toomey, S. (1998). *Communicating effectively with the Chinese*. SAGE Publications.

Gregg, K. R. (1984). Krashen's Monitor and Occam's Razar. *Applied Linguistics*, *5*(2), 79-100. https://doi.org/10.1093/applin/5.2.79

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274. https://doi.org/10.3102/01623737011003255

Ho, D. Y. F. (1976). On the concept of face. *American Journal of Sociology, 81*(4), 867–884. https://doi.org/10.1086/226145

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Liu, H., Lv, C., & Chen, J. (2025). Willingness to communicate and anxiety in human–human and human–chatbot interaction contexts: an idiodynamic investigation. *Innovation in Language Learning and Teaching*, 1–25. https://doi.org/10.1080/17501229.2025.2560101

Macintyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Macintyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: developing an idiodynamic method to capture rapidly changing affect. Applied Linguistics. https://doi.org/10.1093/applin/amq037

Quyen, T. T. (2025). Addressing the “silent classroom”: a theoretical perspective on language anxiety and communicative readiness among vietnamese primary learners. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 4*(4), 10-18. https://doi.org/10.60087/jklst.vol4.n4.002

Sifianou, M. (2011). On the Concept of Face and Politeness. In F. Bargiela-Chiappini & D. Z. Kádár (Eds.), *Politeness Across Cultures* (pp. 42-58). Palgrave Macmillan UK. https://doi.org/10.1057/9780230305939_3

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: a meta-analysis. *Studies in Second Language Acquisition*. https://doi.org/10.1017/s0272263118000311

Khanh, H. D. P., & Ngoc, T. T. N. (2022). EXPLORING VIETNAMESE NON-ENGLISH-MAJORED freshmen's ENGLISH-SPEAKING ANXIETY AT a PUBLIC UNIVERSITY IN VIETNAM. *VNU Journal of Foreign Studies, 38*(5), 105-125. https://doi.org/10.25073/2525-2445/vnufs.4789

Zhou, Z., Li, M., Hodis, G. M., & Croucher, S. M. (2026). Cultural intelligence of native English-speaking teachers in China. International Journal of Intercultural Relations, 111, 102374. https://doi.org/10.1016/j.ijintrel.2026.102374

Zhong, W., Zhou, Z., Wang, L., & Yang, Y. (2026). Developing and validating multi-scenario EFL willingness to communicate scale across classroom, extracurricular, and GenAI-mediated scenarios. BMC Psychology, 14(1), 819. https://doi.org/10.1186/s40359-026-04607-1

#190 BETTER WRITING OR LESS THINKING? AI OVER-RELIANCE IN EFL WRITING CLASSROOMS

Thanh Thuy Nguyen[1]

Abstract

The rapid integration of generative artificial intelligence (AI) tools into English as a Foreign Language (EFL) writing classrooms has intensified debate over whether AI-assisted writing supports or displaces the cognitive processes essential to writing development. While AI tools can facilitate idea generation, grammatical correction, and drafting fluency, concerns remain regarding students' increasing reliance on AI for core writing functions. This study investigates self-reported patterns of perceived AI dependency and their associations with independent writing performance among 113 non-English major undergraduates at a Vietnamese technical university. Adopting a mixed-method design, the study combined survey data with pre–post writing task analysis following a six-week period of AI-assisted writing practice. Writing performance was evaluated across three dimensions: grammatical accuracy, lexical diversity, and idea originality. Findings indicated moderate but pervasive self-reported reliance on AI for idea generation, linguistic decision-making, and evaluative judgment. Although a statistically significant improvement in grammatical accuracy was observed, the effect size was small, and no significant gains emerged in lexical diversity or idea originality. Higher levels of AI reliance were negatively associated with independent writing performance, particularly in relation to idea development. Qualitative findings further suggested that many students approached AI as a habitual productivity tool rather than a metacognitive learning resource. Taken together, the findings point to a productivity–proficiency tension in AI-assisted EFL writing and highlight the need for pedagogically structured AI integration that preserves learners' cognitive engagement in the writing process.

***Keywords*:** AI over-reliance, EFL writing, AI-assisted learning, learner autonomy, cognitive offloading

1. Introduction

1.1 Background of the Study

The emergence of large language model (LLM)-powered writing tools, most prominently ChatGPT, has catalysed a paradigm shift in how students engage with academic writing tasks (Chan & Hu, 2023; Kasneci et al., 2023). Within EFL education, where students simultaneously manage cognitive demands of language learning and content production, these tools offer immediate affordances: on-demand idea generation, instant grammatical correction, and stylistic reformulation. Adoption of generative AI tools has been rapid since the public release of ChatGPT (Tlili et al., 2023; Kasneci et al., 2023).

Yet this rapid uptake has outpaced critical examination of its pedagogical consequences. Emerging evidence suggests a troubling possibility: that tools designed to improve writing output may simultaneously undermine the cognitive processes, including planning, drafting, monitoring, revising, that are essential to writing skill development (Hayes

[1]*Corresponding author, Lecturer, Foreign Language Center, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam; Email: nguyenthanhthuy@hcmut.edu.vn*

& Chenoweth, 2006; Graham & Perin, 2007). This concern is particularly acute in EFL contexts, where writing instruction is understood not merely as a product goal but as a vehicle for linguistic and cognitive development (Manchón, 2011; Hyland, 2019).

1.2 Research Problem

The theoretical construct of cognitive offloading, the externalization of mental operations to environmental resources, provides a useful lens for understanding this tension. While cognitive offloading is not inherently detrimental (Risko & Gilbert, 2016), its systematic application to higher-order writing tasks may reduce the productive struggle associated with deep learning (Sinha & Kapur, 2021; Sweller et al., 2019). When students delegate idea generation, syntactic decisions, and evaluative judgment to AI systems, they may produce better-looking texts while developing shallower cognitive engagement with the writing process itself.

1.3 Research Objectives

This study aims to (1) describe patterns and levels of self-reported AI dependency among EFL undergraduates at a Vietnamese technical university, (2) examine whether a six-week period of AI-assisted writing practice leads to measurable gains in grammatical accuracy, lexical diversity, and idea originality, and (3) investigate the relationship between self-reported AI dependency patterns and independent writing outcomes.

1.4 Research Questions

Against this backdrop, where AI tools promise writing improvement yet may simultaneously erode the cognitive processes that make improvement durable, empirical investigation of actual dependency patterns and their consequences becomes essential. This study addresses this gap through a mixed-method investigation among 113 non-English major undergraduates at Ho Chi Minh City University of Technology (HCMUT), Vietnam. Three research questions guide the inquiry:

RQ1: What are the patterns and extent of perceived AI dependency among EFL undergraduate students in academic writing?

RQ2: To what extent do AI-assisted practice periods produce measurable gains in grammatical accuracy, lexical diversity, and idea originality in student writing?

RQ3: What is the relationship between AI dependency subscale scores and post-task writing performance outcomes?

1.5 Significance of the Study

Despite the rapid adoption of generative AI in EFL writing classrooms, empirical evidence on the pedagogical consequences of over-reliance remains limited, particularly in the Vietnamese higher education context. This study contributes: (1) an internally consistent instrument — AI Dependency in Academic Writing Scale (ADAWS) — for measuring perceived AI dependency across theoretically motivated dimensions; (2) preliminary empirical evidence linking specific dependency patterns to differential writing outcomes; and (3) a process-before-product pedagogical model grounded in the findings. These contributions have practical relevance for EFL educators, curriculum designers, and institutional policymakers navigating the responsible integration of AI tools into academic writing instruction.

2. Literature Review

2.1 AI Tools in EFL Writing Education

Computer-assisted language learning has long incorporated technology into writing instruction, from early grammar-checking and concordancing tools to more recent collaborative digital writing environments (Warschauer & Healey, 1998). Scholarship on AI in EFL writing has shifted considerably in recent years: where earlier work centred on automated writing evaluation (AWE) systems, attention has since turned to generative AI, which can produce full drafts rather than merely evaluate them (Guo & Li, 2024; Lim et al., 2023). Proponents argue that generative AI lowers affective barriers: anxiety reduction, increased drafting fluency, and improved final-product quality have been reported (Song & Song, 2023). In the Vietnamese higher education context, AI-powered tools were perceived to reduce writing anxiety, improve time efficiency, and support idea generation (Thi et al., 2025). However, Deane (2013) argued that genuine writing competence depends on sustained practice in self regulation, planning, and goal setting, all of which may be undermined when learners rely too heavily on AI assistance. Perkins et al. (2023) highlighted challenges associated with identifying and responding to generative-AI-assisted academic writing and potential misuse.

2.2 Cognitive Offloading and Its Consequences for Learning

At moderate levels, the externalisation of mental processes to environmental resources, conceptualised by Risko and Gilbert (2016) as cognitive offloading, represents an adaptive and functionally efficient cognitive strategy. The critical question for education, however, is whether offloading is applied to tasks that are themselves the target of learning. Sweller et al.'s (2019) cognitive load theory provides a foundational framework: writing in a second language is a high-load activity, simultaneously engaging working memory for lexical retrieval, syntactic construction, discourse organization, and content generation. AI tools may reduce extraneous load, but equally may reduce germane load, the productive cognitive effort associated with schema formation and skill acquisition (Paas & Merriënboer, 2020). Sinha and Kapur's (2021) meta-analysis provides evidence that problem solving followed by instruction can support deeper conceptual learning and transfer more effectively than instruction followed by problem solving, particularly when the learning design closely follows the principles of productive failure. Systematic AI delegation may therefore represent a form of premature problem-solving associated with slower long-term skill development. Closely related to these cognitive concerns is the question of whether such delegation also erodes learners' broader capacity for self-directed language learning.

2.3 Learner Autonomy and Self-Regulated Writing

Among the most widely valued outcomes of higher education language programs is learner autonomy: the capacity and disposition to take genuine responsibility for one's own learning (Little, 2015). In writing, autonomy manifests as self-regulated writing (SRW): the ability to plan, monitor, evaluate, and revise without external scaffolding (Graham, 2018; Zimmerman & Moylan, 2009). The relationship between AI use and learner autonomy is theoretically ambiguous. AI tools can scaffold self-regulation by providing feedback learners can selectively incorporate (Hattie & Clarke, 2018). However, when learners habitually accept AI generated output without critical scrutiny, a tendency reflected in the Evaluative Passivity (EVP) subscale, they may fail to cultivate the metacognitive monitoring capacities that underpin autonomous writing proficiency. Shen & Teng (2024) found significant correlations between reduced critical thinking skills and heavier AI-assisted writing use. Taken together, these theoretical concerns underscore the urgent need for empirically grounded tools capable of measuring the specific patterns of AI dependency that may impede autonomous writing development.

2.4 Measuring AI Over-Reliance: Conceptual Gaps

Despite the theoretical concerns outlined above, validated instruments for measuring AI over-reliance in academic writing remain conspicuously absent from the literature. Existing AI literacy scales (Wang et al., 2022; Ng et al., 2021) focus on competence with and attitudes toward AI, not dependency patterns. The present study addresses this gap by developing the AI Dependency in Academic Writing Scale (ADAWS), a 20-item Likert instrument with three theoretically grounded subscales: (1) Idea Generation Reliance (IGR), reflecting dependence on AI for content planning; (2) Linguistic Delegation (LDE), reflecting deference to AI for grammatical, lexical, and syntactic decisions; and (3) Evaluative Passivity (EVP), reflecting failure to critically assess AI output. This framework draws on Manchón's (2011) tripartite model of EFL writing as a cognitive, linguistic, and evaluative activity.

3. Methodology

3.1 Research Context

The study was conducted within the English for Academic Purposes (EAP) writing course at Ho Chi Minh City University of Technology (HCMUT), a leading public technical university in Vietnam. This compulsory course serves undergraduate students across engineering, technology, and science faculties, typically enrolled in Year 1 or Year 2. Class sizes ranged from 25 to 35 students per section. At the time of data collection, HCMUT had no formal institutional policy governing student use of AI writing tools; individual instructors established their own guidelines. Within the study context, students were permitted to use AI tools during out-of-class writing practice activities (Weeks 2–7) but not during timed pre- and post-task assessments conducted under controlled conditions. This setting, in which generative AI had become widely accessible yet pedagogically unstructured, provided a naturalistic context for investigating self-reported AI dependency patterns and their associations with independent writing performance.

3.2 Research Design

A concurrent mixed-method design (Tashakkori & Teddlie, 2010) was adopted, integrating quantitative survey data and pre–post writing performance analysis with open-ended qualitative responses. The quantitative strand addresses RQ1–RQ3 through descriptive statistics, paired-samples t-tests, and Pearson correlation analysis. The qualitative strand, drawn from five open-ended questionnaire items, provides explanatory depth for the quantitative patterns. Integration occurred at the interpretation stage, with qualitative themes used to elaborate quantitative findings (Morse, 2010).

3.3 Participants

The sample comprised 113 undergraduate students enrolled in the English for Academic Purposes (EAP) writing component at HCMUT, all drawn from engineering, technology, and science faculties. The sample comprised 83 males (73.5%), 23 females (20.4%), and 7 who preferred not to specify (6.2%). Year-of-study distribution was approximately balanced across Years 1–4+. All participants had minimum A2-level English proficiency and reported prior use of at least one AI writing tool before the study. Participation was voluntary and responses would not affect course grades. While N = 113 students completed the ADAWS questionnaire, the analytic samples reported in Tables 3–6 range from N = 91 to N = 97 because not every participant completed both the pre- and post-task writing assessments under controlled conditions; absences on the Week 1 or Week 8 testing sessions, and a small number of incomplete or illegible scripts that could not be reliably rated, were the primary sources of attrition. Listwise deletion was applied for each analysis, so the exact N varies slightly depending on which writing dimension and which subscale pairing is being analysed

(e.g., n = 92 for grammatical accuracy, n = 91 for lexical diversity, n = 97 for idea originality). These N values are reported individually in each table note.

3.4 Instruments

3.4.1 AI Over-Reliance Survey

The primary instrument was the ADAWS, a 20-item Likert-scale questionnaire whose items were adapted from Wang et al. (2022) and reoriented toward behavioural dependency in writing contexts. In the present study, perceived AI dependency was operationalised through three dimensions: Idea Generation Reliance (IGR), Linguistic Delegation (LDE), and Evaluative Passivity (EVP). Subscales measure Idea Generation Reliance (8 items; e.g., "I use AI to generate the main ideas of my essays before I start writing"), Linguistic Delegation (7 items; e.g., "I accept AI grammar suggestions without checking them myself"), and Evaluative Passivity (5 items; e.g., "I submit AI-revised drafts without reading them carefully"), rated on a 5-point scale. Present-sample Cronbach's alpha values were .84, .81, and .78 for the three subscales. Importantly, the three-subscale structure is theoretically motivated but has not been empirically confirmed through exploratory factor analysis (EFA) or confirmatory factor analysis (CFA); subscale-level findings should be read as descriptive and hypothesis-generating.

3.4.2 Affective Response Survey

A supplementary nine-item Likert scale, the Affective Response Survey (ARS), assessed emotional and motivational responses to AI-assisted writing (Cronbach's $\alpha = .76$), covering enjoyment, confidence, anxiety, and positive anticipation.

3.4.3 Writing Tasks

Pre- and post-writing tasks consisted of argumentative essay prompts of equivalent difficulty. Students wrote 150-200-word essays within 20 minutes without AI assistance, before and after a six-week AI-assisted practice period. Essays were evaluated using an analytic rubric with three dimensions: grammatical accuracy, lexical diversity, and idea originality (each scored 1–10). A single rater conducted all scoring; inter-rater reliability was not established, which represents a methodological limitation.

3.5 Procedure and Data Analysis

The data collection unfolded across eight weeks, following a structured sequence designed to capture both pre-intervention performance and post-intervention change. In Week 1, participants completed the pre-writing task (no AI access, 20 minutes). In Weeks 2–7, students engaged in AI-assisted writing activities as part of their regular EAP curriculum using tools of their choice. In Week 8, the post-writing task was administered under identical controlled conditions. The questionnaire was administered electronically in Week 7. Descriptive statistics, Cronbach's alpha, paired-samples t-tests (with Cohen's d), and Pearson correlations were computed using Python (SciPy 1.11; pandas 2.0). Qualitative responses underwent inductive thematic analysis (Braun & Clarke, 2019): 47 open codes consolidated into 12 intermediate categories and three overarching themes, reviewed by a second analyst. The following section presents findings from both strands, beginning with participant characteristics and scale performance, before moving to writing outcomes and their associations with AI dependency.

3.6 Reliability Analysis

Internal consistency of the ADAWS subscales was evaluated using Cronbach's alpha prior to main analyses. All values met or exceeded the threshold of .80 recommended for research instruments (George & Mallery, 2003): Idea Generation Reliance (IGR): α = .947 (excellent); Linguistic Delegation (LDE): α = .895 (good); Evaluative Passivity (EVP): α = .859 (good); and the Affective Response Survey (ARS): α = .936 (excellent). These values provide confidence in the reliability of subscale scores for the subsequent analyses. Inter-item correlations were reviewed to confirm no redundant or misfit items. It should be noted that while alpha values are strong, the three-subscale structure has not been confirmed through exploratory or confirmatory factor analysis; subscale-level findings should therefore be treated as descriptive and hypothesis-generating pending dedicated psychometric validation.

3.7 Ethical Considerations

The study was conducted in accordance with standard ethical principles for educational research. Participation was voluntary, and students were explicitly informed prior to data collection that their responses would have no bearing on course grades or academic standing. Electronic informed consent was obtained before questionnaire administration. All data were anonymised at the point of collection and stored securely on password-protected institutional servers accessible only to the researcher. Participants were free to withdraw at any point without penalty. The study involved no deception, no collection of sensitive personal data, and no intervention beyond normal coursework activities.

4. Results

4.1 Participant Characteristics

An overview of the sample's demographic profile is presented in Table 1. ChatGPT was the most frequently used AI writing tool (61.9%), followed by Grammarly (15.9%) and Google Gemini (15.0%). Notably, 48.7% of participants reported daily AI use for writing, and a further 27.4% reported use several times per week, indicating that AI-assisted writing is already a habitual practice for the majority of this cohort.

Table 1

Participant Demographic Profile (N = 113)

Variable	Category	n	%
Gender	Male	83	73.5%
	Female	23	20.4%
	Prefer not to say	7	6.2%
Year of Study	Year 1	31	27.4%
	Year 2	30	26.5%
	Year 3	24	21.2%
	Year 4+	23	20.4%
Primary AI Tool	ChatGPT	70	61.9%

	Grammarly	18	15.9%
	Google Gemini	17	15.0%
Usage Frequency	Daily	55	48.7%
	Several times/week	31	27.4%
	Occasionally	22	19.5%
	Rarely	5	4.4%

4.2 ADAWS: Internal Consistency and Descriptive Statistics

Reliability estimates across all scales were strong, with Cronbach's alpha values of .947 (IGR), .895 (LDE), .859 (EVP), and .936 (ARS). Inter-subscale correlations were computed as a preliminary check: IGR–LDE r = .71, IGR–EVP r = .64, LDE–EVP r = .58 (all $p < .001$). These moderate-to-high correlations suggest that the three dimensions may not be fully orthogonal, cautioning against interpreting subscale-specific effects as evidence of distinct constructs pending formal EFA/CFA. Variance inflation factors (VIFs) ranged from 2.83 to 3.12, indicating multicollinearity did not reach levels that typically invalidate ordinary least squares (OLS) estimates.

All three subscale means clustered closely around the scale midpoint: IGR M = 2.99 (SD = 0.73), LDE M = 3.02 (SD = 0.61), and EVP M = 3.09 (SD = 0.60), indicating moderate reliance across all dimensions. At the item level, EVP2 ("I trust AI evaluations of my writing quality more than my own judgment") recorded the highest mean (M = 3.16, SD = 0.79). One-way ANOVA revealed that daily AI users scored significantly higher on IGR compared to occasional users [$F(2, 110) = 4.21$, $p = .017$]. With reliance patterns thus established, we turn to the question of whether these patterns correspond to measurable changes in writing performance over the study period.

Table 2

ADAWS Item-Level and Subscale Descriptive Statistics and Internal Reliability (N = 113)

Item	Statement (abbreviated)	M	SD	α	Subscale M (SD)
IGR1	Use AI to generate main ideas	3.13	0.77		
IGR2	Ask AI for topic suggestions	3.00	0.85		
IGR3	Rely on AI for thesis statements	2.84	0.90		
IGR4	Use AI-generated outlines	3.00	0.80		
IGR5	Cannot begin essay without AI	2.89	0.85		
IGR6	Use AI to decide supporting points	2.97	0.96		
IGR7	Difficulty generating ideas without AI	3.04	0.91		

IGR8	Essays reflect AI thinking more than own	3.05	0.81	.947	2.99 (0.73)
LDE1	Accept AI grammar suggestions w/o checking	2.96	0.86		
LDE2	Rely on AI to choose vocabulary	3.00	0.79		
LDE3	Use AI to correct all grammar errors	3.08	0.85		
LDE4	Don't worry about grammar — AI will fix it	3.11	0.77		
LDE5	Let AI decide sentence structure	2.97	0.78		
LDE6	Use AI phrases w/o evaluating fit	3.05	0.69		
LDE7	Language shaped primarily by AI output	3.00	0.66	.895	3.02 (0.61)
EVP1	Submit AI-revised drafts without reading	3.11	0.83		
EVP2	Trust AI evaluations over own judgment	3.16	0.79		
EVP3	Apply AI feedback without questioning	3.08	0.75		
EVP4	Rarely assess if AI content matches meaning	3.02	0.77		
EVP5	Consider essay finished when AI approves	3.05	0.74	.859	3.09 (0.60)

*Note. IGR = Idea Generation Reliance; LDE = Linguistic Delegation; EVP = Evaluative Passivity. Subscale α reported in the final item row of each subscale.

4.3 Pre- and Post-Writing Performance

Examining the three writing dimensions in turn, grammatical accuracy showed a statistically significant improvement [$t(91) = -4.51$, $p < .001$], with post-task scores ($M = 6.11$, $SD = 1.73$) exceeding pre-task scores ($M = 5.75$, $SD = 1.25$). The effect size was small (Cohen's $d = 0.24$) — below the threshold of what is typically considered educationally meaningful in L2 writing research. Examining individual performance, 68.5% of participants showed some grammar gain, while 31.5% showed no improvement or a decline. For lexical diversity, the pre-to-post difference was negligible and non-significant [$t(90) = -1.12$, $p = .265$, $d = 0.08$]. Idea originality improvement was similarly non-significant [$t(96) = -1.76$, $p = .081$, $d = 0.10$]. Together, the six-week AI-assisted practice period was associated with meaningful improvement only in surface-level grammatical accuracy, with no detectable gains in deeper dimensions within this timeframe. The next analysis examines whether these limited performance outcomes are systematically related to the degree of AI dependency each student exhibited.

Table 3

Pre- and Post-Task Writing Performance: Paired-Samples t-Test Results (N = 91–97)

Criterion	Pre-M	Pre-SD	Post-M	Post-SD	t	p	d
Grammatical Accuracy	5.75	1.25	6.11	1.73	−4.51	< .001	0.24
Lexical Diversity	5.65	0.93	5.74	1.19	−1.12	.265	0.08
Idea Originality	5.30	1.43	5.45	1.65	−1.76	.081	0.10

*Note. Scores are on a 1–10 analytic rubric scale. Effect sizes reported as Cohen's d. N varies across dimensions due to missing or incomplete pre- and post-task writing scripts.

Table 4

Hierarchical Regression of Pre-Task Grammar and ADAWS Subscales Predicting Post-Task Grammatical Accuracy (N = 91)

Predictor	B	SE	β	t	p	ΔR²
Model 1						
Pre-task grammatical accuracy	0.412	—	—	< .001		.175
Model 2						
Pre-task grammatical accuracy	0.389	—	—	< .001		
IGR (Idea Generation Reliance)	−0.934	—	−0.501	< .001		
LDE (Linguistic Delegation)	−1.042	—	−0.523	< .001		
EVP (Evaluative Passivity)	−0.135	—	−0.072	.381		.243

*Note. Model 1: Pre-task grammatical accuracy only ($R^2 = .175$, $p < .001$). Model 2: Pre-task grammar + IGR + LDE + EVP ($R^2 = .418$, $\Delta R^2 = .243$, $p < .001$). IGR ($\beta = -0.501$, $p < .001$) and LDE ($\beta = -0.523$, $p < .001$) remained significant negative predictors after controlling for baseline proficiency; EVP did not ($\beta = -0.072$, $p = .381$). B = unstandardized coefficient; SE = standard error; β = standardized coefficient. Dash (—) indicates SE not separately reported for Model 1 baseline predictor. N varies across dimensions due to missing or incomplete pre- and post-task writing scripts.

To examine whether the negative associations between AI dependency and writing performance could be accounted for by pre-existing proficiency differences, a hierarchical regression analysis was conducted (Table 4). Because the IGR, LDE, and EVP subscales have not yet been confirmed through exploratory or confirmatory factor analysis (see Section 3.4.1), they are entered into the models below as theoretically motivated composite scores rather than as psychometrically validated latent constructs; regression coefficients should accordingly be interpreted as indicating that these item clusters are statistically associated with post-task performance, not as evidence that IGR, LDE, and EVP are independent, validated constructs. Model 1 entered pre-task grammatical accuracy as a baseline covariate; Model 2 additionally entered the three ADAWS subscales. Model 1 explained 17.5% of variance in post-task

grammar scores ($R^2 = .175$, $p < .001$). Adding the ADAWS subscales produced a significant incremental improvement ($\Delta R^2 = .243$, $p < .001$), raising total explained variance to $R^2 = .418$. Even after controlling for baseline proficiency, both IGR ($\beta = -0.501$, $p < .001$) and LDE ($\beta = -0.523$, $p < .001$) remained significant negative predictors of post-task grammar, while EVP did not reach significance ($\beta = -0.072$, $p = .381$). These findings indicate that self-reported AI dependency patterns remain statistically associated with post-task grammatical performance after adjustment for baseline proficiency. While causal inferences cannot be drawn from the present design, the results suggest that the observed relationships are not solely attributable to differences in initial writing proficiency.

4.4 Relationships Between AI Dependency and Writing Performance

Turning to the associations between dependency patterns and writing outcomes, both IGR ($r = -.46$, $p < .001$) and LDE ($r = -.43$, $p < .001$) showed significant negative correlations with post-task grammatical accuracy, while EVP yielded a smaller but still significant negative relationship ($r = -.30$, $p = .003$). Post-task absolute scores for lexical diversity and idea originality showed no significant subscale associations. However, gain score analysis revealed a more differentiated picture: IGR was significantly and negatively associated with idea originality gains ($r = -.45$, $p < .001$, $n = 97$), with LDE ($r = -.24$, $p = .020$) and EVP ($r = -.23$, $p = .025$) also predicting smaller gains. This pattern suggests that the null result for post-task idea scores reflects distributional compression in absolute scores rather than a genuine absence of dependency effects.

A proficiency-stratified analysis adds important nuance. Students with higher baseline proficiency (pre-grammar ≥ median = 5.56) showed larger grammar gains ($M = 0.53$, $SD = 0.81$) than lower-proficiency peers ($M = 0.20$, $SD = 0.70$; $t(90) = 2.07$, $p = .042$), and reported lower Linguistic Delegation ($M = 2.84$ vs. 3.17; $t(111) = -2.97$, $p = .004$). This is consistent with a reverse causality account: lower-proficiency students both rely more heavily on AI and are less equipped to leverage AI feedback productively.

Table 5

Pearson Correlations Between ADAWS Subscales and Post-Task Writing Performance Scores, with Idea Originality Gain-Score Correlations (N = 91–97)

Subscale	Post-Grammar r	p	Post-Lex r	p	Post-Idea r	p	Gain-Idea r	p
IGR (Idea Generation Reliance)	−.460	< .001	−.031	.756	.041	.681	−.450	< .001
LDE (Linguistic Delegation)	−.426	< .001	.127	.202	.177	.070	−.240	.020
EVP (Evaluative Passivity)	−.296	.003	−.091	.363	−.129	.189	−.230	.025

**Note. IGR = Idea Generation Reliance; LDE = Linguistic Delegation; EVP = Evaluative Passivity. Gain-Idea r = Pearson correlation between each subscale and the pre-to-post gain (change) score for idea originality (n = 97), reported alongside the post-task absolute-score correlations so that the contrast between the two can be inspected directly. The near-zero post-task absolute correlations (Post-Idea r ranging from .041 to −.129, all p > .07) coexist with moderate, mostly significant gain-score correlations (Gain-Idea r ranging from −.230 to −.450), indicating that post-task absolute idea-originality scores were compressed within a narrow range across participants regardless of dependency level, while the amount of change from pre- to post-task was*

more clearly patterned by AI dependency. This is the distributional compression referred to in Section 4.4. N varies across dimensions due to missing or incomplete pre- and post-task writing scripts.

A simultaneous multiple regression analysis with IGR, LDE, and EVP predicting post-task grammar was statistically significant [$F(3, 87) = 11.43$, $p < .001$], explaining approximately 28% of the variance ($R^2 = .283$, adjusted $R^2 = .258$). Both IGR ($\beta = -0.624$, $p < .001$) and LDE ($\beta = -0.658$, $p < .001$) emerged as significant negative predictors; EVP did not reach significance after controlling for the other subscales ($\beta = -0.098$, $p = .312$). A supplementary model including pre-task grammatical accuracy as a baseline covariate yielded $R^2 = .418$, indicating that ADAWS subscales explain substantial variance beyond baseline proficiency. Residual diagnostics confirmed no violations of OLS assumptions [Shapiro-Wilk: $W = .984$, $p = .328$; Breusch-Pagan: $\chi^2(3) = 4.17$, $p = .243$].

Table 6

Multiple Regression of ADAWS Subscales Predicting Post-Task Grammatical Accuracy (N = 91)

Predictor	B	SE	β	t	p	sr^2
IGR (Idea Generation Reliance)	−1.167	0.271	−0.624	−4.31	< .001	.132
LDE (Linguistic Delegation)	−1.318	0.307	−0.658	−4.29	< .001	.128
EVP (Evaluative Passivity)	−0.275	0.272	−0.098	−1.01	.312	.007

**Note. $R^2 = .283$, adjusted $R^2 = .258$; $F(3, 87) = 11.43$, $p < .001$. B = unstandardized coefficient; SE = standard error; β = standardized coefficient; sr^2 = squared semi-partial correlation. Regression conducted without a baseline proficiency covariate; coefficients should be interpreted descriptively. *p < .001. N varies across dimensions due to missing or incomplete pre- and post-task writing scripts.*

4.5 Affective Responses

Students' emotional responses to AI-assisted writing were generally positive, with the ARS returning strong internal consistency ($\alpha = .936$) and subscale means ($M = 3.01$, $SD = 0.67$) indicating moderately favourable affect overall. ARS1 ("I find AI-assisted writing tasks enjoyable") returned $M = 3.10$ ($SD = 0.95$). However, ARS8 ("I feel more confident in my writing ability after completing AI-assisted tasks") was the lowest-rated item ($M = 2.88$, $SD = 0.84$), with 34.5% rating it "Disagree" or "Strongly Disagree." This dissociation, moderate enjoyment paired with low confidence in independent ability, is consistent with the performance data showing limited independent skill gains, and aligns with Krashen and Mason's (2020) observation that a low-anxiety environment may be necessary but is not sufficient for acquisition.

4.6 Qualitative Themes

4.6.1. Theme 1: AI as a Writing Crutch

Surfacing across approximately 71% of responses, the most pervasive theme was students' own recognition that their AI use had become habitual rather than strategic. A representative student wrote: "I know I use ChatGPT too much. When I have a writing task, my first move is to open it. I don't even try to think first anymore." Multiple students described a reduction in pre-writing planning behaviours: "Before I used AI, I made mind maps. Now I don't bother."

4.6.2. Theme 2: Confidence Without Competence

Approximately 38% of respondents articulated awareness of a gap between AI-assisted and unaided writing quality: "My AI essays look good. But when I write in the exam without it, I struggle with ideas and grammar. I know the gap is getting bigger." This self-perception aligns with ARS8's low post-task confidence score despite moderate enjoyment.

4.6.3. Theme 3: Pragmatic Motivation Over Learning Intent

Students' motivations for AI use were predominantly pragmatic, including saving time, meeting deadlines, and producing higher quality submissions, rather than being oriented toward learning objectives, accounting for approximately 64% of responses. Only approximately 18% framed AI use as a deliberate learning strategy. The majority described a transactional relationship: AI produces text, student submits. Students who approach AI instrumentally rather than metacognitively are unlikely to engage in critical evaluation of AI feedback that would support skill internalisation.

5. Discussion

5.1 Summary of Findings by Research Question

To bring the quantitative and qualitative strands together coherently, the findings are first summarised against each of the three research questions before being discussed thematically in the sections that follow. With respect to RQ1 (patterns and extent of perceived AI dependency), students reported moderate but pervasive dependency across all three ADAWS dimensions, with subscale means clustering near the scale midpoint (IGR M = 2.99, LDE M = 3.02, EVP M = 3.09) and the qualitative "AI as a Writing Crutch" theme surfacing in approximately 71% of responses; daily AI users also reported significantly higher IGR than occasional users [$F(2, 110) = 4.21$, $p = .017$]. With respect to RQ2 (whether a six-week AI-assisted practice period produces measurable writing gains), the evidence was mixed and modest: grammatical accuracy improved significantly but with a small effect size ($d = 0.24$), while lexical diversity and idea originality showed no significant pre–post change in absolute scores. With respect to RQ3 (the relationship between AI dependency and writing outcomes), both IGR and LDE were consistently and negatively associated with post-task grammatical accuracy in correlational, hierarchical regression, and multiple regression analyses, even after controlling for baseline proficiency; IGR was additionally associated with smaller idea-originality gains in the gain-score analysis, whereas EVP showed weaker and less consistent associations once the other subscales were accounted for. Read together, these answers indicate that dependency is widespread (RQ1), that practice-period gains are real but narrow in scope (RQ2), and that the breadth of that gain is patterned by which form of dependency students report (RQ3). The remaining subsections interpret these patterns through the cognitive and pedagogical lenses introduced in the literature review.

5.2 The Productivity–Proficiency Tension

Collectively, the quantitative and qualitative findings converge on a coherent picture of AI-mediated writing practice in this context. At its core, what emerges is a productivity–proficiency tension: AI tools are associated with surface-level improvement in students' output, yet this does not appear to coincide with gains in independent writing competencies. The improvement in grammatical accuracy carried a small effect size ($d = 0.24$), falling below the range of effect sizes reported across writing-instruction interventions in Graham and Perin's (2007) meta-analysis. This modest grammar gain must also be contextualised by the absence of a control group: comparable improvement may occur in EFL courses without AI

involvement. Students who self-reported the heaviest perceived reliance on AI for idea generation and linguistic decisions scored lowest on independent grammatical production, a moderate negative association ($r \approx -.43$ to $-.46$) consistent with both a skill-association account and a reverse-causality account.

Gain score analysis revealed that IGR was negatively associated with idea originality gains ($r = -.45$, $p < .001$), comparable in magnitude to the grammar correlations. Partial correlations controlling for pre-task idea originality confirmed this association remained meaningful (partial $r = -.31$, $p = .003$). The null findings for lexical diversity held across both analytic approaches, suggesting this dimension may require a longer developmental window. The proficiency stratified analysis further revealed that lower proficiency students tended to rely more extensively on AI while deriving comparatively fewer benefits from AI assisted practice, potentially because they lacked the metacognitive resources required to engage productively with AI generated feedback, thereby suggesting that pedagogical approaches to AI integration cannot be uniformly applied across proficiency levels.

These patterns resonate with evidence that heavy AI use in writing is associated with reduced self-directed learning competency and diminished critical thinking (Shen & Teng, 2024), and with broader concerns about whether technology-mediated writing assistance produces durable skill gains (Perkins et al., 2023). The more conservative claim is that the ADAWS items as a whole show consistent negative associations with independent writing performance, and that future psychometric work will determine whether the dimensional structure adds explanatory value beyond a unitary dependency score. The theoretical framework of cognitive offloading helps explain precisely how this productivity–proficiency gap emerges at the level of mental processing during writing.

5.3 Cognitive Offloading and the Writing Process

A coherent account of these findings can be found in the theoretical framework of cognitive offloading (Risko & Gilbert, 2016). Writing in a second language simultaneously activates orthographic, lexical, syntactic, discourse-level, and rhetorical processes (Hayes, 2012; Kellogg et al., 2013). When AI systems handle idea generation, grammatical monitoring, and evaluative judgment, the cognitive load associated with these processes may be substantially reduced, but so too is the practice that makes them automatic and internalisable. Sweller et al.'s (2019) cognitive load theory predicts precisely this outcome: extraneous load reduction is beneficial only when germane cognitive engagement is preserved. When AI handles the germane tasks, learning opportunities may be reduced.

Vocabulary development requires repeated, contextually rich, generative encounters with words (Nation, 2022; Schmitt, 2010), while task-induced involvement can influence both incidental vocabulary acquisition and subsequent retention (Gao et al., 2024). Thus, passive exposure to AI-generated text may not provide sufficient opportunities for the active processing and engagement needed for durable lexical learning. Similarly, original idea generation requires the activation and recombination of prior knowledge (Weisberg, 2014), and habitual AI use for ideation was associated with lower idea-originality gains in this sample. The qualitative data offer corroboration: students reporting habitual AI use for idea generation also reported diminished pre-writing planning behaviours. This pattern of reduced generative engagement may not remain confined to ideation; it extends to the metacognitive and evaluative dimensions of the writing process, as the following section discusses.

5.4 Evaluative Passivity as a Metacognitive Risk

Perhaps the most telling single finding concerns EVP2: "I trust AI evaluations of my writing quality more than my own judgment", which returned the highest mean across the entire scale, pointing to a worrying displacement of metacognitive self-assessment by algorithmic authority. Self-evaluation is the foundation of self-regulated writing development (Graham, 2018; Zimmerman & Moylan, 2009). EVP's significant negative correlation with post-task grammar ($r = -.30$, $p = .003$), though the smallest of the three subscales, is consistent with the interpretation that the evaluative dimension of AI dependency carries independent performance-related variance. Students in the "Confidence Without Competence" theme explicitly recognized a widening gap between AI-assisted and unaided performance, yet this awareness had not translated into altered AI use behaviours, suggesting that metacognitive awareness alone is insufficient without structured pedagogical scaffolding. These insights directly inform the pedagogical model proposed below, which aims to preserve the cognitive and evaluative processes that AI delegation tends to displace.

5.5 Implications for EFL Pedagogy

When considered together, the findings raise serious questions about the wisdom of allowing unrestricted AI access at every stage of the drafting process. The formative function of writing assignments, as practice for autonomous cognitive engagement, may be constrained when students routinely delegate planning, drafting, and evaluation to AI systems. A "process-before-product" integration model is suggested by the data, illustrated through a four-step classroom sequence:

Step 1 (10 minutes): Students complete a timed AI-free brainstorm in response to the writing prompt, generating their own ideas and a rough outline before any AI interaction occurs.

Step 2 (15–20 minutes): Students write a full first draft without AI, preserving the ideational and syntactic effort that builds internal language schemas.

Step 3 (10 minutes): Students use AI exclusively for targeted grammar feedback on specific structures identified as developmental priorities, not for content generation or full-text rewriting.

Step 4 (5 minutes): Students submit a brief reflection log identifying one grammatical pattern from the AI feedback, explaining the error and how they intend to avoid it in future writing.

This sequence preserves generative cognitive processes within the learner while leveraging AI's strengths as a targeted feedback tool, consistent with Hattie and Clarke's (2018) framework of effective feedback. The qualitative findings further underscore the need for explicit metacognitive instruction concerning the ways in which AI use shapes students' cognitive engagement and long term development, without which learners are unlikely to regulate their AI use in pedagogically productive ways (Veenman et al., 2006; Pintrich, 2004).

5.6 Limitations

Any interpretation of the present findings should be tempered by several methodological constraints worth making explicit. First, the sample was drawn from a single university with a predominantly male, STEM-oriented cohort; findings should not be generalised to EFL learners more broadly. The gender imbalance, reflecting structural features of technology faculty enrolment in Vietnamese higher education, means findings may not fully represent the AI dependency patterns of female EFL writers. Second, the six-week practice period is a significant constraint. Writing development in a second language is longitudinal;

gains in lexical diversity and idea originality may require months or years to manifest. The null findings in these dimensions should not be interpreted as evidence of no effect, but rather as an artefact of insufficient developmental time. Third, the absence of a randomised control condition reflects a deliberate methodological choice prioritising ecological validity. Single-rater scoring without inter-rater reliability data also represents a limitation. Fourth, the sample encompasses both grammar-assistance tools (Grammarly, n = 18) and generative AI platforms (ChatGPT and Gemini combined, n = 87) — tools that differ fundamentally in cognitive demands. Preliminary comparisons found generative AI users reported significantly higher IGR (M = 3.07 vs. 2.63; t(103) = 2.84, p = .005) and LDE (M = 3.09 vs. 2.71; t(103) = 2.61, p = .010), suggesting that dependency patterns are concentrated in the generative AI sub-group and future studies should treat tool type as a substantive moderator. Fifth, because ADAWS is a self-report instrument, it captures students' perceived behavioural dependency tendencies rather than objectively observed AI use behaviours. Future studies should triangulate self-report data with behavioural measures such as prompt logs, AI interaction histories, or screen recordings to establish convergent validity. Sixth, the construct validity of ADAWS requires dedicated future work: the three-subscale structure is theoretically grounded but has not been tested through EFA or CFA, and all subscale-level findings should be read accordingly.

6. Conclusion

Beginning with a concern about the pedagogical consequences of habitual AI use, the inquiry traced patterns of dependency and their associations with independent writing performance among non-English major undergraduates at a single Vietnamese technical university. Through the development and deployment of ADAWS and pre–post writing performance comparison, a consistent pattern has emerged within this sample: moderate but pervasive AI dependency coexists with negligible independent writing skill gains across six weeks of AI-assisted practice. Because the sample was drawn from one institution, with a strong STEM orientation and a pronounced gender skew (73.5% male), these patterns should be read as specific to this tech-oriented cohort rather than as general claims about EFL learners; whether they extend to humanities-oriented, more gender-balanced, or non-Vietnamese EFL populations is an empirical question for future research, not an assumption of the present study.

Three contributions to the emerging literature are worth highlighting. First, it pilots an exploratory instrument, an internally consistent instrument for measuring behavioural AI dependency across three theoretically motivated dimensions, addressing the absence of validated dependency-specific measures in the field. Second, it provides preliminary evidence that AI dependency patterns may be differentially associated with writing outcomes: gain score analysis shows IGR is associated with reduced growth in idea originality ($r = -.45$, $p < .001$), while proficiency-stratified analysis shows lower-proficiency students report higher linguistic delegation yet smaller performance gains, pointing toward a moderating role of baseline ability. Third, the findings suggest a potential relationship between students' perceptions of AI dependency, affective responses to AI-assisted writing, and writing outcomes. Although students generally reported moderately positive affect toward AI-assisted writing, qualitative findings revealed concerns regarding confidence in independent writing performance, indicating that affective and motivational factors may play an important role in shaping the educational consequences of AI use.

Together, these findings move the conversation from whether AI affects writing to which dimensions of AI dependence are associated with writing outcomes, for whom, and in what circumstances. These associations should not be read as evidence that AI dependency causes poorer writing performance: the cross-sectional, non-randomised design used here

cannot rule out reverse causality (students with weaker independent skills may turn to AI more heavily) or the influence of unmeasured third variables, and the absence of a control group means part of the observed grammar gain may simply reflect normal classroom progress unrelated to AI use. What the data do support, within this single-institution, predominantly male, STEM-oriented sample, is a consistent statistical association between self-reported dependency patterns and writing performance that future controlled and longitudinal research should test directly. Integration, as the data suggest, is not inherently beneficial: the mode, timing, and pedagogical framing of AI use may determine whether tools serve as scaffolds toward autonomy or substitutes for the productive cognitive engagement that writing development requires, though this study's design permits this only as a plausible interpretation rather than a demonstrated mechanism. Future research should seek to develop and empirically validate structured frameworks for AI integration, including process before product models, metacognitive AI literacy curricula, and formative AI feedback protocols, that capitalise on the affordances of generative AI while preserving the indispensable cognitive processes involved in learning to write. Nevertheless, the interpretation of the present findings should be approached with caution in light of the methodological limitations discussed earlier in this study.

References

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, 43. https://doi.org/10.1186/s41239-023-00411-8

Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. *Assessing Writing, 18*(1), 7–24. https://doi.org/10.1016/j.asw.2012.10.002

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.

Gao, M., Qian, J., & Rasool, U. (2024). Effects of task-induced involvement and time on task on incidental L2 vocabulary acquisition. *SAGE Open, 14*(2). https://doi.org/10.1177/21582440241249340

Graham, S. (2018). A revised writer(s)-within-community-of-writers model of writing. *Educational Psychologist, 53*(4), 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System, 124*, 103362. https://doi.org/10.1016/j.system.2024.103362

Hattie, J., & Clarke, S. (2018). *Visible learning: Feedback*. Routledge.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication, 23*(2), 135–149. https://doi.org/10.1177/0741088306286283

Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Stadler, M. (2023). Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: A progress report. *Journal of Writing Research, 5*(2), 159–190. https://doi.org/10.17239/jowr-2013.05.02.1

Krashen, S., & Mason, B. (2020). The optimal input hypothesis: Not all comprehensible input is of equal value. *CATESOL Newsletter*, 5(1), 1–2.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: ragnarök or reformation? a paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Little, D. (2015). Language learner autonomy: Rethinking the goal. In M. Everhard & L. Murphy (Eds.), *Assessment and autonomy in language learning* (pp. 12–28). Palgrave Macmillan.

Manchón, R. M. (2011). *Learning-to-write and writing-to-learn in an additional language*. John Benjamins.

Morse, J. M. (2010). Procedures and practice of mixed method design: Maintaining control, rigor, and complexity. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavioral research* (2nd ed., pp. 339–352). SAGE.

Nation, I. S. P. (2022). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/9781009093873

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Paas, F., & Merriënboer, J. J. G. V. (2020). Cognitive-load theory: methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science, 29*(4), 394-398. https://doi.org/10.1177/0963721420922183

Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2023). Detection of GPT-4 generated text in higher education: combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics, 22*(1), 89-113. https://doi.org/10.1007/s10805-023-09492-6

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Shen, X., & Teng, M. F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity, 52*, 101524. https://doi.org/10.1016/j.tsc.2024.101524

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research, 91*(5), 761–798. https://doi.org/10.3102/00346543211019105

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Sweller, J., Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tashakkori, A., & Teddlie, C. (2010). Putting the human back in "human research methodology": The researcher in mixed methods research. *Journal of Mixed Methods Research, 4*(4), 271–277. https://doi.org/10.1177/1558689810382532

Thi, X. H. N., Thien, H. V. H., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through AI-powered tools: Perceived benefits and challenges among Vietnamese EFL students. *Discover Education, 4*, 472. https://doi.org/10.1007/s44217-025-00905-9

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: chatgpt as a case study of using chatbots in education. *Smart Learning Environments, 10*(1). https://doi.org/10.1186/s40561-023-00237-x

Veenman, M. V. J., Hout-Wolters, B. H. A. M. V., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3-14. https://doi.org/10.1007/s11409-006-6893-0

Wang, B., Rau, P. L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology, 42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31*(2), 57–71. https://doi.org/10.1017/S0261444800012970

Weisberg, R. W. (2014). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning, 21*(1), 5-39. https://doi.org/10.1080/13546783.2014.886625

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge.

Bionote

Thanh Thuy Nguyen

nguyenthanhthuy@hcmut.edu.vn

Serving as a lecturer at Ho Chi Minh University of Technology (HCMUT) since 2008, Thanh Thuy Nguyen has accumulated over 15 years of experience in English language teaching. She is committed to enhancing students' language proficiency through engaging lessons, interactive activities, and meaningful extracurricular programs that foster holistic development. She holds a Bachelor's degree in English Language Teaching from Ho Chi Minh City University of Pedagogy and a Master's degree in Applied Linguistics from Curtin University, Australia. Guided by her philosophy that "teaching is a dual learning process," she believes teachers and students grow together through continuous learning and reflection.

#193 TEACHER BELIEFS, IDENTITY, AND PROFESSIONAL AGENCY IN ESL-ORIENTED TOURISM EDUCATION: A PLS-SEM ANALYSIS IN VIETNAM

Phong Hai Tran[1]

Abstract

The transition from English as a Foreign Language (EFL) to English as a Second Language (ESL)-oriented pedagogy has become increasingly critical in tourism education, particularly in emerging economies such as Vietnam where English proficiency directly shapes graduate employability and service quality. Drawing on teacher cognition theory, sociocultural perspectives, and professional agency theory, this study develops and empirically tests a structural model linking teacher beliefs, professional identity, and professional agency to pedagogical innovation and teacher-perceived communicative competence in tourism English education. Using survey data collected from 112 English lecturers in tourism and hospitality programs across universities in Ho Chi Minh City, Vietnam, the study employs Partial Least Squares Structural Equation Modeling (PLS-SEM) to examine both linear and indirect relationships. The findings reveal that teacher beliefs regarding communicative and industry-relevant English significantly influence identity reconstruction and agency enactment, which in turn drive pedagogical innovation (e.g., experiential learning, VR/AR integration) and enhance teacher-perceived communicative competence. Professional agency emerges as a key mediating mechanism translating beliefs into practice under institutional constraints. This study contributes to the literature by integrating teacher cognition with ESL transition in tourism education and provides actionable implications for curriculum reform, teacher development, and policy in Vietnam and comparable contexts.

Keywords: teacher beliefs, professional agency, pedagogical innovation, ESL-oriented tourism education, teacher-perceived communicative competence

1. Introduction

The globalization of tourism and the increasing mobility of international travelers have positioned English as a critical lingua franca in hospitality and tourism industries, thereby reshaping the objectives of language education in non-English-speaking countries. In this context, tourism education must move beyond traditional EFL paradigms toward more practice-oriented, ESL-like environments where English is used as a tool for authentic communication rather than as an abstract academic subject (Cloudia Ho, 2020). This transformation is particularly salient in Vietnam, where tourism has become a strategic economic sector and universities are under pressure to produce graduates with strong communicative competence and workplace readiness.

[1] *Corresponding author, Ho Chi Minh City University of Foreign Languages-Information Technology, Ho Chi Minh City, Vietnam; Email: phongth@huflit.edu.vn*

Despite this demand, prior research consistently highlights a mismatch between tourism English training and industry expectations, particularly in developing contexts where language instruction remains grammar-focused and examination-driven(Al-Tarawneh & Osam, 2019; Ghany & Latif, 2012) . This gap suggests that structural reform alone is insufficient; rather, understanding teachers as key agents of change becomes essential. Teachers' beliefs, identity, and professional agency fundamentally shape how curricular innovations are interpreted and implemented in practice.

While previous studies have examined communicative language teaching (CLT) and technology-enhanced learning in tourism education (Elfiondri et al., 2021; He et al., 2025), limited research integrates these pedagogical developments with teacher cognition and agency frameworks (Pantaruk et al., 2026), particularly using robust empirical methods such as PLS-SEM. Moreover, there is a lack of Vietnam-specific empirical evidence examining how teachers navigate the transition toward ESL-oriented tourism education (Şimşek & Kalıpçı, 2023).

1.1 Problem Statement

Although Vietnam's tourism sector increasingly requires graduates who can communicate fluently and appropriately in authentic service encounters, tourism English instruction at the tertiary level remains predominantly grammar-oriented, examination-driven, and detached from workplace communication demands. Reform efforts have concentrated on revising curricula and adopting new technologies, yet these structural interventions have produced limited improvement in graduates' communicative readiness. The central problem this study addresses is that the success of the transition from an EFL to an ESL orientation depends less on curricular documents than on whether teachers—who ultimately interpret and enact reform—hold the beliefs, reconstruct the professional identities, and exercise the agency required to translate ESL-oriented intentions into innovative classroom practice. At present, it remains unclear which teacher-level mechanisms actually convert reform intentions into pedagogical innovation and, in turn, into improved communicative outcomes, leaving institutions without an evidence-based basis for prioritizing their limited development resources.

1.2 Research Gap

Three gaps motivate this study. First, although communicative language teaching (CLT) and technology-enhanced learning have been examined extensively in tourism education (Elfiondri et al., 2021; He et al., 2025), few studies integrate these pedagogical developments with teacher cognition and professional agency frameworks within a single, empirically testable model; existing work tends to treat beliefs, identity, agency, and innovation in isolation rather than specifying the sequential mechanism that links them. Second, prior research frequently assumes that reconstructing teacher identity is sufficient to generate innovative practice, yet the distinct mediating role of professional agency in converting identity into action remains empirically underexamined. Third, despite growing policy attention to English-medium employability in Vietnam, there is a shortage of Vietnam-specific, model-based evidence on how teachers navigate the transition toward ESL-oriented tourism education (Şimşek & Kalıpçı, 2023). Addressing these gaps requires a robust quantitative approach capable of testing an integrated mediation model, which this study provides through PLS-SEM.

1.3 Scope and Delimitations

This study is deliberately bounded in several respects. Empirically, it focuses on English lecturers teaching tourism and hospitality programs at universities in Ho Chi Minh

City, Vietnam—a leading tourism and educational hub—and does not claim to represent all Vietnamese regions or institutions. Conceptually, the model is delimited to five teacher-centered constructs (teacher beliefs, professional identity, professional agency, pedagogical innovation, and teacher-perceived communicative competence) and intentionally excludes institutional-level antecedents such as leadership, workload, and policy constraints, as well as student-generated performance data, which lie beyond the study's teacher-cognition focus. The outcome construct is operationalized as teachers' perceptions of student communicative competence rather than as objective or student-reported proficiency; accordingly, the findings should be interpreted as reflecting instructor-perceived outcomes. These boundaries define the analytical scope of the study and are revisited as limitations and directions for future research in Section 8.

1.4 Research Objectives and Questions

To address the problem and gaps outlined above, this study develops and empirically tests a structural model linking teacher beliefs, professional identity, and professional agency to pedagogical innovation and teacher-perceived communicative competence in tourism English education. Specifically, it seeks to answer three research questions: (RQ1) How do teacher beliefs shape professional identity and professional agency in the transition toward ESL-oriented tourism education? (RQ2) Through what mechanism do professional identity and professional agency translate into pedagogical innovation? and (RQ3) To what extent does pedagogical innovation enhance teacher-perceived communicative competence? By answering these questions, the study clarifies the teacher-level mechanisms underpinning ESL-oriented pedagogical transformation and identifies where institutional resources can be most effectively directed.

2. Literature Review and Hypothesis Development

2.1 Theoretical Framework

The conceptual model of this study integrates three complementary theoretical lenses—teacher cognition theory, sociocultural theory, and professional agency theory—to explain how the transition from an EFL to an ESL orientation is enacted in tourism English education. Teacher cognition theory provides the cognitive foundation of the model, positing that teachers' beliefs, knowledge, and assumptions about language and learning function as the primary filter through which curricular reforms are interpreted and instructional decisions are made (Borg, 2003; Wyatt & Dikilitıaş, 2019). Within this framework, teacher beliefs are treated as the antecedent that initiates the transformation process.

Sociocultural theory complements this cognitive foundation by locating teacher development within social, institutional, and professional contexts. From this perspective, professional identity is not a fixed attribute but a dynamic, socially negotiated construct that evolves as teachers interact with institutional expectations, industry demands, and emerging pedagogical roles (Langum & Sullivan, 2020; Tao & Gao, 2017). Sociocultural theory therefore explains how beliefs are mediated by context to reshape teachers' self-perceptions in ESL-oriented settings, moving them from knowledge transmitters toward facilitators and designers of experiential learning.

Professional agency theory supplies the behavioral bridge between cognition and practice. Agency—teachers' capacity to make intentional pedagogical choices and act purposefully within, and at times against, institutional constraints—is the mechanism that converts beliefs and reconstructed identities into concrete instructional innovation (Priestley et al., 2015; Tao & Gao, 2017). Integrating these three lenses yields a coherent cognitive–

behavioral logic in which cognition (beliefs) is filtered through socially constructed identity and enacted through agency to produce pedagogical innovation and, ultimately, teacher-perceived communicative competence. Figure 1 operationalizes this integrated framework, and the hypotheses developed in the following subsections (2.2–2.5) trace each pathway in this sequence.

2.2 Teacher Beliefs and ESL-Oriented Pedagogy

Teacher beliefs represent deeply held assumptions about teaching and learning that guide instructional practices (Wyatt & Dikilitaş, 2019). In tourism education, beliefs about communicative competence, industry relevance, and experiential learning are particularly critical. Communicative Language Teaching (CLT) has been shown to significantly enhance students' communicative competence and confidence in tourism English contexts (Liu et al., 2024; Şimşek & Kalıpçı, 2023). However, the adoption of CLT and similar approaches depends largely on teachers' epistemological orientations and perceived teaching goals. In emerging ESL-like environments, teachers who believe in the importance of authentic communication and workplace alignment are more likely to adopt experiential and student-centered approaches. Conversely, teachers who maintain traditional beliefs may resist pedagogical change. Therefore:

H1: Teacher beliefs positively influence professional identity reconstruction toward ESL-oriented teaching.

H2: Teacher beliefs positively influence professional agency in tourism English education.

2.3 Professional Identity Transformation

Teacher identity is not static but evolves through interaction with institutional contexts, professional expectations, and pedagogical demands (Langum & Sullivan, 2020). In tourism education, the shift toward ESL requires teachers to move from knowledge transmitters to facilitators of experiential learning and industry-relevant communication (Ju et al., 2025). The integration of immersive technologies such as VR and AR has further expanded the teacher's role as a designer of learning environments that simulate real-world communication scenarios, thereby enhancing student engagement and reducing anxiety (Akyürek et al., 2024). Identity reconstruction thus becomes a crucial mechanism through which beliefs are translated into pedagogical practices. Therefore:

H3: Professional identity positively influences professional agency.

H4: Professional identity positively influences pedagogical innovation.

2.4 Professional Agency as a Mediator

Professional Agency refers to teachers' capacity to act purposefully and make pedagogical decisions within institutional constraints (Habib et al., 2026). In tourism education, agency enables teachers to adapt curricula, integrate technology, and align teaching with industry needs (Liu et al., 2024). Studies on AR, VR, and experiential learning demonstrate that innovative teaching practices significantly improve student engagement and practical competencies (Ju et al., 2025). However, the adoption of such practices depends on teachers' ability to exercise agency.

H5: Professional agency positively influences pedagogical innovation.

2.5 Pedagogical Innovation and Student Outcomes

Pedagogical innovation in tourism education increasingly encompasses a range of learner-centered and practice-oriented approaches, notably experiential learning, technology-enhanced instruction such as virtual reality (VR) and augmented reality (AR), and industry-based simulations that replicate authentic workplace scenarios (Ju et al., 2025). These approaches shift the focus from traditional knowledge transmission toward immersive and interactive learning environments in which students actively engage in communication tasks reflective of real tourism contexts. Empirical evidence suggests that experiential learning and simulation-based activities enable students to develop practical language skills, enhance confidence, and bridge the gap between classroom knowledge and industry expectations, while technology-enhanced tools such as VR and AR further enrich learning by providing realistic, low-risk environments for practicing communication and decision-making (Akyürek et al., 2024; Rahmanu & Molnár, 2026). In addition, multimedia resources such as video-based instruction and digital platforms contribute to vocabulary acquisition and contextual understanding, thereby strengthening communicative competence (Cloudia Ho, 2020). Collectively, these pedagogical innovations not only improve student engagement but also foster employability skills by aligning language learning with professional practice in tourism settings (Kacetl & Frydrychova-Klimova, 2015). Accordingly, it is hypothesized that pedagogical innovation has a significant positive effect on teacher-perceived communicative competence.

H6: Pedagogical innovation positively influences teacher-perceived communicative competence.

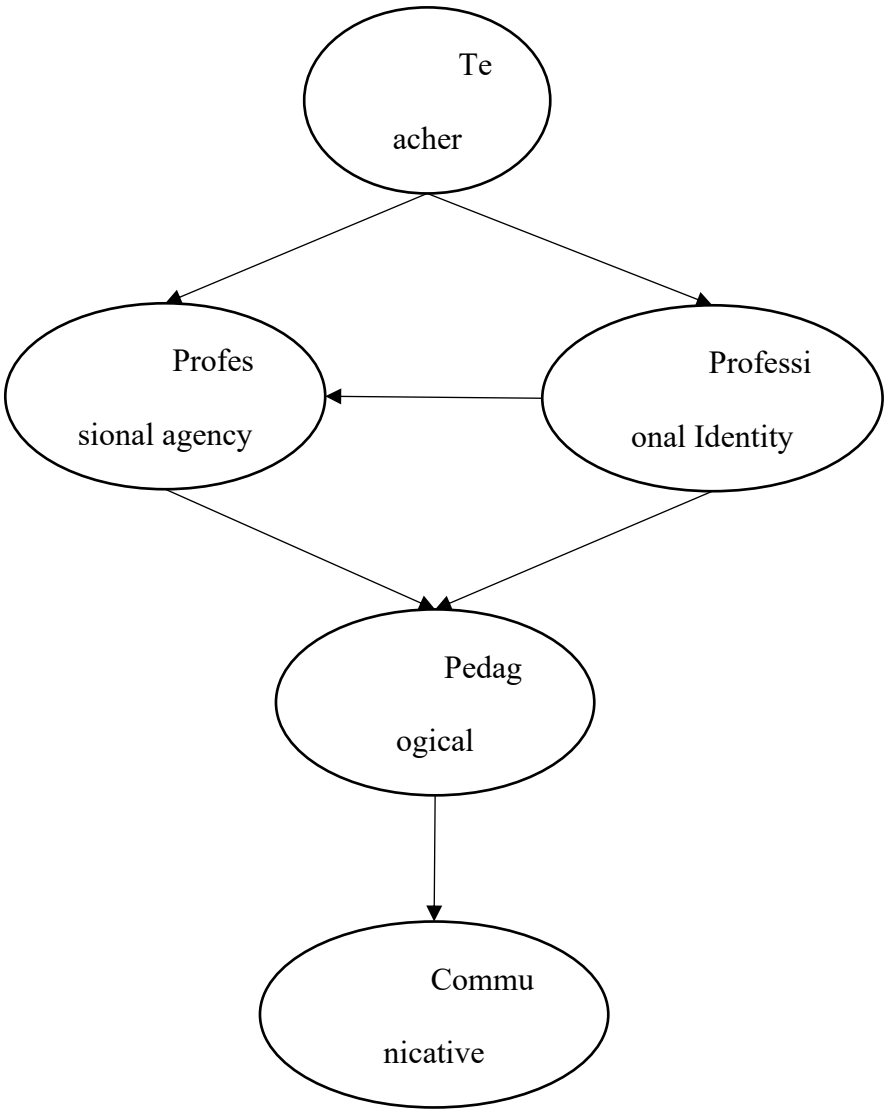
3. Research Model

The proposed research model, presented in Figure 1, is grounded in an integrated theoretical framework that combines teacher cognition, sociocultural theory, and professional agency perspectives to explain pedagogical transformation in ESL-oriented tourism education. Specifically, the model comprises five key constructs: Teacher Beliefs (TB), Professional Identity (PI), Professional Agency (PA), Pedagogical Innovation (PIN), and Teacher-Perceived Communicative Competence (CC). Teacher beliefs (TB) refer to instructors' underlying assumptions about language teaching, particularly the importance of communicative competence and industry-relevant English in tourism contexts, which have been shown to shape instructional practices and learning outcomes (Cloudia Ho, 2020). These beliefs influence the development of professional identity (PI), defined as teachers' evolving self-perception as facilitators of communication, designers of experiential learning. In tourism education, identity reconstruction is increasingly necessary as teaching shifts toward more authentic, practice-based approaches supported by immersive technologies such as VR and AR (Takada et al., 2022). Professional agency (PA) represents teachers' capacity to make pedagogical decisions, adapt curricula, and implement innovative teaching strategies within institutional constraints, playing a critical mediating role in translating beliefs and identity into actual classroom practices. This agency is particularly important for adopting pedagogical innovation (PIN), which includes experiential learning, simulation-based instruction, and technology-enhanced approaches that have been demonstrated to improve student engagement and practical competencies in tourism education (Rahmanu & Molnár, 2026; Takada et al., 2022). Finally, teacher-perceived communicative competence (CC) reflects teachers' assessment of students' ability to use English effectively in real-world tourism situations, a core outcome of tourism English programs and a key determinant of employability (Al-Tarawneh & Osam, 2019; Ghany & Latif, 2012). Collectively, these five constructs form a

coherent structural model that captures both the cognitive and behavioral mechanisms underpinning the transition toward ESL-oriented pedagogy in tourism education.

Figure 1

Research model (CC = Teacher-Perceived Communicative Competence)



4. Methodology

Data were collected from 112 English lecturers teaching tourism and hospitality programs at universities located in Ho Chi Minh City, Vietnam, which serves as a major educational and tourism hub with strong industry–academic linkages. Respondents were required to have at least one year of experience teaching Tourism English or English for Specific Purposes (ESP) courses to ensure adequate familiarity with industry-oriented language instruction. A structured questionnaire was administered using both online and offline distribution methods to maximize response rates and capture a diverse sample of instructors across public and private institutions. All measurement items were assessed using a seven-

point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), allowing for greater sensitivity and variance in responses, which is particularly suitable for PLS-SEM analysis.

The measurement scales used in this study were adapted and contextualized from established literature in communicative language teaching (CLT), English for Specific Purposes (ESP), and tourism education to ensure both content validity and contextual relevance. Specifically, Teacher Beliefs (TB) were operationalized to capture instructors' orientations toward communicative, industry-relevant, and practice-based language teaching, consistent with prior studies emphasizing the importance of communicative competence and workplace alignment in tourism English education (Cloudia Ho, 2020; Ghany & Latif, 2012). The TB construct included four items: TB1 ("I believe English teaching should focus on real-world communication"), TB2 ("Industry-relevant English is more important than grammar accuracy"), TB3 ("Students should learn English through interaction and practice"), and TB4 ("Tourism English should reflect workplace communication needs"). Professional Identity (PI) was measured to reflect teachers' evolving roles in ESL-oriented environments, particularly their shift toward facilitation, experiential design, and industry alignment. This construct aligns with research highlighting the transformation of teacher roles in tourism education and technology-enhanced learning contexts (Rahmanu & Molnár, 2026; Takada et al., 2022). Four items were used: PI1 ("I see myself as a facilitator of communication"), PI2 ("I integrate industry practices into my teaching"), PI3 ("My role includes preparing students for real-world tourism contexts"), and PI4 ("I design learning experiences rather than only deliver content"). Professional Agency (PA) was operationalized to capture teachers' capacity to make pedagogical decisions, innovate, and adapt to contextual demands, consistent with prior studies emphasizing the importance of agency in implementing educational innovations (Tao & Gao, 2017). The scale included PA1 ("I adapt teaching methods based on student needs"), PA2 ("I experiment with new teaching approaches"), PA3 ("I integrate technology into my teaching"), and PA4 ("I modify curriculum content to align with industry demands"). Pedagogical Innovation (PIN) was measured to reflect the extent to which teachers adopt experiential, simulation-based, and technology-enhanced teaching practices, which have been widely recognized as critical for improving engagement and practical competencies in tourism education (Ju et al., 2025). The items included PIN1 ("I use experiential learning activities"), PIN2 ("I incorporate simulations or role-plays"), PIN3 ("I use digital tools such as VR, video, or AR in teaching"), and PIN4 ("I design interactive classroom activities"). Finally, Teacher-Perceived Communicative Competence (CC) was measured based on teachers' perceptions of student learning outcomes in tourism contexts. Because all CC items capture instructors' perceptions rather than objective or student-reported proficiency, the construct is labeled Teacher-Perceived Communicative Competence throughout the study to reflect its measurement basis accurately, reflecting the central goal of ESL-oriented instruction. This construct is consistent with prior research emphasizing the importance of communicative ability, confidence, and real-world interaction skills in tourism English education (Cloudia Ho, 2020). The scale included CC1 ("Students can communicate effectively in tourism contexts"), CC2 ("Students demonstrate confidence in speaking English"), CC3 ("Students can handle real-life service situations"), and CC4 ("Students show improved interaction skills").

Data analysis was conducted using Partial Least Squares Structural Equation Modeling (PLS-SEM) with Smart-PLS 4, following a two-stage approach that included assessment of the measurement model and the structural model, in line with established guidelines for variance-based SEM (Hair et al., 2011; Henseler et al., 2014). PLS-SEM was selected due to its suitability for exploratory research, prediction-oriented models, and complex relationships involving mediation effects, as well as its robustness with relatively small to medium sample

sizes. To determine the minimum required sample size, this study employed G*Power, a widely used tool for statistical power analysis in behavioral and social sciences research. Following the guidelines for structural equation modeling, the sample size was calculated based on the maximum number of predictors pointing to an endogenous construct in the proposed model. In this study, the most complex endogenous construct is Pedagogical Innovation (PIN), which has two predictors (Professional Identity and Professional Agency). Accordingly, an a priori power analysis was conducted using the F-test: Linear multiple regression (fixed model, R^2 deviation from zero) with the following parameters: a medium effect size ($f^2 = 0.15$), a significance level (α) of 0.05, and a statistical power ($1 - \beta$) of 0.80, which are commonly recommended thresholds in SEM research (Hair et al., 2023). The results from G*Power indicated that the minimum required sample size is 68 observations. Given that this study collected data from 112 respondents, the sample size exceeds the minimum requirement, thereby ensuring adequate statistical power for detecting significant relationships within the model. This also aligns with the "10-times rule" in PLS-SEM, further supporting the adequacy and robustness of the sample size for model estimation and hypothesis testing.

5. Results and Data Analysis

The demographic profile of the respondents is presented in Table 1. Among the 112 English lecturers surveyed in Ho Chi Minh City, the majority were female (58.9%), while male respondents accounted for 41.1%. In terms of age distribution, most participants were between 30 and 39 years old (46.4%), followed by those under 30 (25.0%), indicating a relatively young and dynamic teaching workforce. Regarding teaching experience, a substantial proportion of respondents had between four and six years of experience (32.1%), while 26.8% had seven to ten years, suggesting that most participants possessed moderate to high levels of professional experience. In terms of educational qualifications, the majority held a master's degree (64.3%), followed by doctoral degrees (19.6%), reflecting a well-qualified academic sample. Additionally, 60.7% of respondents were affiliated with public universities, while 39.3% worked in private institutions, ensuring representation across different institutional contexts. With respect to experience in teaching Tourism English or ESP courses, most lecturers reported more than four years of experience, highlighting their familiarity with industry-oriented language instruction. Overall, the sample demonstrates a balanced and experienced group of educators suitable for examining ESL-oriented pedagogical transformation in tourism education.

Table 1

Demographic Profile of Respondents (N = 112)

Characteristics	Category	Frequency (n)	Percentage (%)
Gender	Male	46	41.1
	Female	66	58.9
Age	Under 30	28	25.0
	30–39	52	46.4
	40–49	22	19.6
	50 and above	10	8.9

Teaching Experience	1–3 years	24	21.4
	4–6 years	36	32.1
	7–10 years	30	26.8
	More than 10 years	22	19.6
Education Level	Bachelor's Degree	18	16.1
	Master's Degree	72	64.3
	Doctoral Degree	22	19.6
Type of Institution	Public University	68	60.7
	Private University	44	39.3
ESP/Tourism English Experience	1–3 years	30	26.8
	4–6 years	40	35.7
	More than 6 years	42	37.5

Sources: Author

5.1 Measurement Model

Table 2 presents the results of the measurement model assessment, including indicator loadings, internal consistency reliability, convergent validity, and multicollinearity diagnostics for all constructs. Overall, the results demonstrate that the measurement model satisfies the recommended criteria for reliability and validity in PLS-SEM analysis (Hair et al., 2011). First, the indicator loadings for all items exceed the acceptable threshold of 0.70, with most values ranging from 0.79 to 0.90, indicating strong item reliability. Although one item (CC4 = 0.681) is slightly below the ideal threshold, it remains acceptable as it is close to 0.70 and does not adversely affect the overall construct reliability. This suggests that all indicators adequately represent their respective latent constructs. Second, internal consistency reliability is confirmed through Cronbach's alpha and composite reliability (ρa and ρc). All constructs exhibit Cronbach's alpha values above 0.80, ranging from 0.816 (CC) to 0.898 (TB), exceeding the recommended threshold of 0.70. Similarly, composite reliability values (ρc) range from 0.880 to 0.929, indicating a high level of internal consistency across all constructs. Third, convergent validity is established through the Average Variance Extracted (AVE). All constructs demonstrate AVE values above the recommended minimum of 0.50, ranging from 0.648 (CC) to 0.766 (TB), confirming that each construct explains more than half of the variance of its indicators. Finally, multicollinearity was assessed using Variance Inflation Factor (VIF) values. All VIF values are below the conservative threshold of 3.0, indicating that there are no significant collinearity issues among the indicators (Ooi & Tan, 2016).

Table 2

Indicator Loadings, Internal Consistency Reliability, Convergent Validity, and Multicollinearity Diagnostics for All Constructs

Construct	Items	Loadings	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)	VIF
CC	**CC1**	0.829	0.816	0.820	0.880	0.648	2.157
	CC2	0.820					2.086
	CC3	0.877					2.675
	CC4	0.681					1.240
PA	**PA1**	0.819	0.881	0.881	0.918	0.737	1.913
	PA2	0.875					2.599
	PA3	0.863					2.781
	PA4	0.875					2.619
PI	**PI1**	0.834	0.871	0.873	0.912	0.721	2.194
	PI2	0.803					1.711
	PI3	0.890					2.861
	PI4	0.868					2.369
PIN	**PIN1**	0.845	0.840	0.874	0.889	0.668	1.811
	PIN2	0.857					1.924
	PIN3	0.790					2.269
	PIN4	0.776					2.230
TB	**TB1**	0.861	0.898	0.904	0.929	0.766	2.387
	TB2	0.844					2.157
	TB3	0.892					2.829
	TB4	0.903					2.966

Table 3 presents the assessment of discriminant validity using the Heterotrait–Monotrait ratio (HTMT), following the recommended threshold of 0.90 for conceptually similar constructs. The HTMT values range from 0.531 to 0.860, all of which are below the conservative cut-off value of 0.90, thereby confirming adequate discriminant validity among the constructs (Ab Hamid et al., 2017). Specifically, the highest HTMT value is observed between Professional Agency (PA) and Professional Identity (PI) (HTMT = 0.860), which is expected given the conceptual proximity of these constructs within teacher cognition and agency frameworks. Similarly, the relationship between Teacher Beliefs (TB) and Professional Identity (PI) shows a relatively high HTMT value (0.839), reflecting their theoretical interdependence. However, both values remain within acceptable limits, indicating that the constructs are empirically distinct despite their conceptual relatedness. Lower HTMT values are observed between Pedagogical Innovation (PIN) and other constructs, such as PIN–TB (0.531) and PIN–PI (0.594), suggesting clearer differentiation between pedagogical practices and underlying cognitive or identity-related constructs. Additionally, the HTMT values involving Communicative Competence (CC) are all below 0.75, further supporting the

distinctiveness of outcome variables from antecedent constructs. Overall, the HTMT results provide strong evidence of discriminant validity, confirming that all constructs in the model are sufficiently distinct and capturing different theoretical dimensions. This supports the robustness of the measurement model and justifies proceeding with structural model evaluation.

Table 3

Discriminant Validity-HTMT Results

	CC	PA	PI	PIN	TB
CC					
PA	0.743				
PI	0.699	0.860			
PIN	0.631	0.747	0.594		
TB	0.737	0.777	0.839	0.531	

5.2 Structural Model Assessment

Table 4 presents the results of the structural model analysis, including path coefficients, t-values, p-values, and confidence intervals obtained through bootstrapping procedures. Overall, the findings provide strong empirical support for most of the hypothesized relationships, highlighting the critical role of teacher cognition and agency in driving pedagogical innovation and communicative outcomes. First, Teacher Beliefs (TB) exhibit a strong and statistically significant positive effect on Professional Identity (PI) ($\beta = 0.753$, $t = 11.094$, $p < 0.001$), supporting H1. This indicates that teachers who hold strong beliefs in communicative and industry-oriented language teaching are more likely to reconstruct their professional identity toward facilitative and experiential roles. Similarly, TB also has a significant positive effect on Professional Agency (PA) ($\beta = 0.292$, $t = 3.382$, $p = 0.001$), supporting H2, suggesting that teacher beliefs contribute directly to their capacity to act and make pedagogical decisions. Second, Professional Identity (PI) significantly influences Professional Agency (PA) ($\beta = 0.539$, $t = 6.451$, $p < 0.001$), supporting H3. This finding highlights that identity reconstruction plays a key role in enabling teachers to exercise agency and adapt their teaching practices. However, the direct relationship between Professional Identity and Pedagogical Innovation (PIN) is not statistically significant ($\beta = 0.092$, $t = 0.963$, $p = 0.336$), leading to the rejection of H4. This suggests that identity alone may not be sufficient to drive innovation unless it is translated into action through agency. In contrast, Professional Agency (PA) demonstrates a strong and significant positive effect on Pedagogical Innovation (PIN) ($\beta = 0.603$, $t = 6.657$, $p < 0.001$), supporting H5. This finding confirms that agency is a critical mechanism through which teachers implement innovative practices such as experiential learning and technology integration. Finally, Pedagogical Innovation (PIN) significantly influences Communicative Competence (CC) ($\beta = 0.564$, $t = 7.786$, $p < 0.001$), supporting H6, indicating that innovative teaching approaches substantially enhance students' ability to communicate effectively in tourism contexts. Overall, the results suggest a sequential mechanism in which teacher beliefs shape identity and agency, while professional agency plays a central role in driving pedagogical innovation, which ultimately leads to improved communicative competence. The non-significant relationship between professional identity and pedagogical innovation further emphasizes the importance of agency as a mediating factor in translating cognitive and identity-related constructs into practical teaching outcomes.

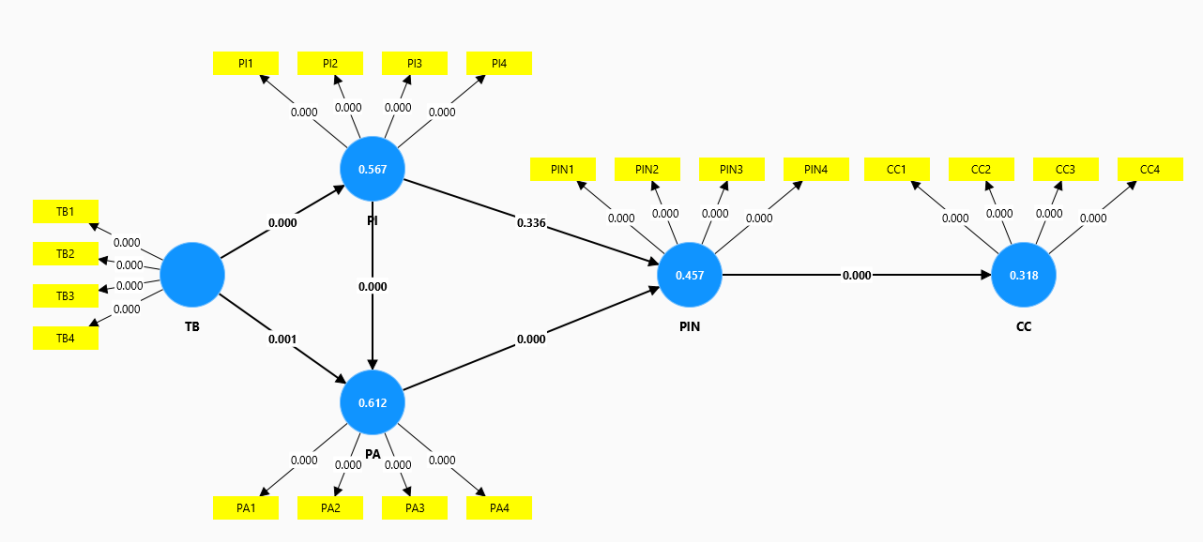
Table 4

Structural Model Results

Hypothesis	Path	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values	2.50%	97.50%	Remark
H1	**TB -> PI**	0.753	0.746	0.068	11.094	0.000	0.592	0.857	Supported
H2	**TB -> PA**	0.292	0.285	0.086	3.382	0.001	0.115	0.456	Supported
H3	**PI -> PA**	0.539	0.543	0.083	6.451	0.000	0.374	0.708	Supported
H4	**PI -> PIN**	0.092	0.091	0.096	0.963	0.336	-0.095	0.278	Unsupported
H5	**PA -> PIN**	0.603	0.602	0.091	6.657	0.000	0.419	0.776	Supported
H6	**PIN -> CC**	0.564	0.568	0.072	7.786	0.000	0.414	0.697	Supported

Figure 2

Structural Model Assessment



Tables 5, 6, and 7 present the results of the structural model evaluation, including explanatory power (R^2), predictive relevance (Q^2), effect sizes (f^2), and out-of-sample predictive performance using PLS-Predict (Shmueli et al., 2019). Collectively, these results provide robust evidence of the model's explanatory and predictive capabilities. As shown in Table 5, the coefficient of determination (R^2) indicates that the model demonstrates moderate to substantial explanatory power across all endogenous constructs. Specifically, Professional Agency (PA) exhibits the highest explanatory power ($R^2 = 0.612$), followed by Professional

Identity (PI) (R^2 = 0.567), Pedagogical Innovation (PIN) (R^2 = 0.457), and Communicative Competence (CC) (R^2 = 0.318). These values suggest that the model explains a considerable proportion of variance, particularly for PA and PI, which are central mediating constructs. Furthermore, predictive relevance (Q^2) values, calculated using the blindfolding procedure, are all greater than zero (CC = 0.189, PA = 0.432, PI = 0.390, PIN = 0.280), confirming that the model has satisfactory predictive relevance for all endogenous constructs (Giovanis et al., 2021; Kock, 2015).

Table 5

R-square and Q^2 (=1-SSE/SSO)

	R-square	R-square adjusted		SSO	SSE	Q^2 (=1-SSE/SSO)
CC	0.318	0.312	**CC**	456	369.903	0.189
PA	0.612	0.605	**PA**	456	259.171	0.432
PI	0.567	0.563	**PI**	456	278.080	0.390
PIN	0.457	0.447	**PIN**	456	328.213	0.280

Table 6 reports the effect sizes (f^2) to assess the relative impact of exogenous constructs on endogenous variables. The results indicate that Teacher Beliefs (TB) have a very large effect on Professional Identity (PI) (f^2 = 1.310), highlighting the dominant role of beliefs in shaping identity reconstruction. Additionally, Professional Identity has a moderate effect on Professional Agency (PA) (f^2 = 0.323), while Professional Agency demonstrates a substantial effect on Pedagogical Innovation (PIN) (f^2 = 0.285). Pedagogical Innovation also shows a strong effect on Communicative Competence (CC) (f^2 = 0.465), confirming its critical role in enhancing student outcomes. In contrast, the effect of Professional Identity on Pedagogical Innovation is negligible (f^2 = 0.007), further supporting the non-significant path observed in the structural model and reinforcing the importance of agency as the key driver of innovation.

Table 6

Effect size

	CC	PA	PI	PIN	TB
CC					
PA				0.285	
PI		0.323		0.007	
PIN	0.465				
TB		0.095	1.310		

Table 7 presents the results of the PLS-Predict analysis, which evaluates the model's out-of-sample predictive performance by comparing PLS-SEM predictions with those from a linear regression model (LM) and a naïve benchmark (IA). The Q^2predict values are mostly positive across indicators, indicating that the model has predictive relevance at the indicator level. Moreover, for the majority of items, the PLS-SEM model produces lower prediction errors (RMSE and MAE) compared to the benchmark models, suggesting superior predictive

performance. Although a few indicators (e.g., PIN4) show slightly weaker predictive accuracy, the overall pattern demonstrates that the model achieves medium to high predictive power.

Table 7

PLS predicts

	Q²predict	PLS-SEM_RMSE	PLS-SEM_MAE	LM_RMSE	LM_MAE	IA_RMSE	IA_MAE
CC1	0.157	1.169	0.985	1.135	0.926	1.273	1.045
CC2	0.217	1.197	0.944	1.241	0.788	1.353	1.031
CC3	0.210	1.199	0.987	1.292	0.889	1.349	1.113
CC4	0.064	1.528	1.264	1.561	1.293	1.580	1.326
PA1	0.487	0.934	0.753	0.970	0.730	1.304	0.997
PA2	0.264	1.247	1.000	1.267	1.026	1.453	1.186
PA3	0.248	1.321	1.041	1.356	1.067	1.524	1.242
PA4	0.347	1.097	0.880	1.117	0.895	1.358	1.109
PI1	0.283	1.299	1.001	1.310	1.005	1.534	1.269
PI2	0.546	0.881	0.662	0.895	0.639	1.308	0.842
PI3	0.344	1.174	0.844	1.195	0.884	1.449	1.156
PI4	0.371	1.187	0.938	1.226	0.953	1.497	1.234
PIN1	0.310	1.180	0.997	1.185	0.983	1.421	1.132
PIN2	0.232	1.228	1.003	1.246	1.022	1.402	1.157
PIN3	0.027	1.354	1.044	1.389	1.078	1.372	1.053
PIN4	-0.012	1.539	1.243	1.546	1.251	1.529	1.236

Table 8 and Figure 3 report the results of the Importance–Performance Map Analysis (IPMA), which extends the PLS-SEM findings by simultaneously considering the total effects (importance) of antecedent constructs on the target variable (Communicative Competence, CC) and their average latent variable scores (performance). This analysis provides actionable insights by identifying which constructs have the greatest impact on CC and where managerial or pedagogical improvements should be prioritized. The results indicate that Pedagogical Innovation (PIN) has the highest importance (effect = 0.564) but relatively low performance (57.797), suggesting that although it is the most influential driver of communicative competence, its current implementation level remains insufficient. This highlights PIN as the primary area for improvement, implying that greater emphasis should be placed on enhancing experiential learning, simulation-based activities, and technology integration to maximize student outcomes. In contrast, Teacher Beliefs (TB) demonstrate moderate importance (0.276) but the highest performance (82.992), indicating that lecturers already possess strong beliefs aligned with communicative and industry-oriented teaching. Similarly, Professional Identity (PI) shows relatively high performance (78.310) with moderate importance (0.235), suggesting that teachers generally perceive themselves as facilitators of communication and experiential learning. However, their direct contribution to communicative competence is comparatively

weaker. Professional Agency (PA) exhibits both moderate importance (0.340) and relatively high performance (74.526), indicating that teachers are reasonably active in adapting teaching methods and implementing innovations. Given its mediating role in the structural model, PA remains a crucial lever for translating beliefs and identity into effective pedagogical practices.

When compared to the mean values (importance = 0.354; performance = 73.406), the IPMA results clearly identify Pedagogical Innovation as the key priority for strategic improvement, while Teacher Beliefs and Professional Identity can be considered strengths of the current system. These findings provide practical guidance for policymakers and educational institutions, emphasizing the need to invest in innovative teaching practices and technological integration to enhance communicative competence in tourism education.

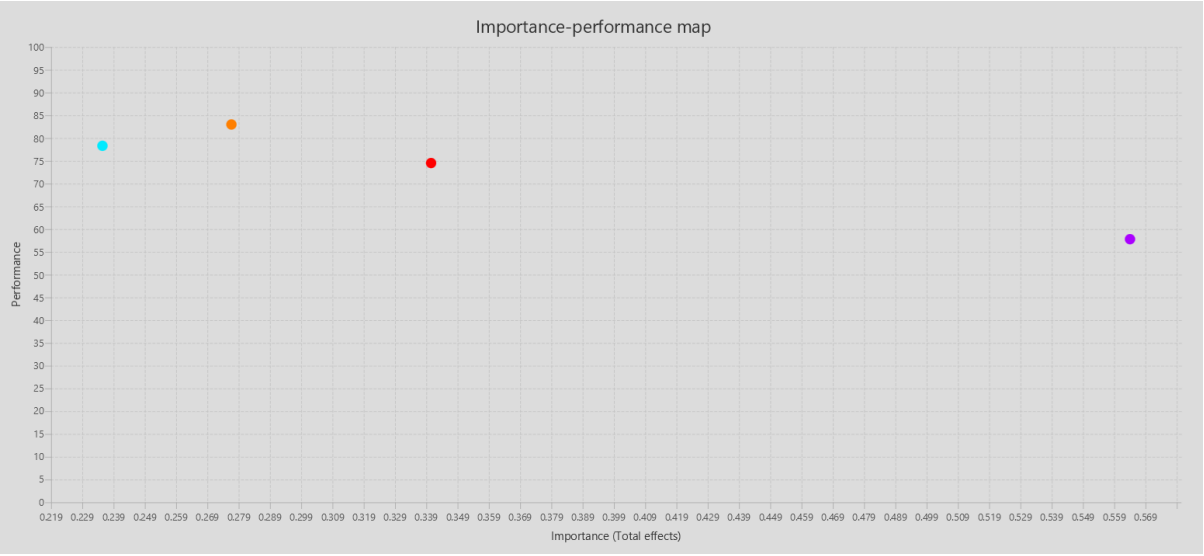
Table 8

Importance-performance analysis

	Effect	Performance
PA	0.340	74.526
PI	0.235	78.310
PIN	0.564	57.797
TB	0.276	82.992
Mean	**0.354**	**73.406**

Figure 3

Importance–Performance Map Analysis



6. Discussion

This study set out to explain how teacher beliefs, professional identity, and professional agency jointly shape pedagogical innovation and, ultimately, teacher-perceived communicative competence in ESL-oriented tourism education. The results provide a coherent, sequential

mechanism such as beliefs → identity → agency → innovation → outcomes with several important nuances that extend prior work.

First, the strong and positive effect of teacher beliefs on professional identity (H1) and professional agency (H2) confirms that what teachers believe about language learning, particularly the value of communicative, industry-relevant English, acts as a primary driver of how they see themselves and how they act in classrooms. This finding is consistent with evidence from communicative language teaching (CLT), which shows that teachers who prioritize authentic communication and interaction are more likely to adopt learner-centered practices and pursue workplace alignment (Cloudia Ho, 2020). It also aligns with ESP-oriented research in tourism education indicating that industry relevance and real-world communication needs reshape instructional orientations (Ghany & Latif, 2012). In the Vietnamese context, where traditional, grammar-focused instruction still coexists with emerging communicative demands, these results underscore that belief change is a necessary starting point for pedagogical transformation.

Second, the significant effect of professional identity on professional agency (H3) highlights identity reconstruction as a key enabling condition for action. Teachers who perceive themselves as facilitators, designers of learning experiences, and mediators between academia and industry are more likely to exercise agency, adapting methods, integrating tools, and revising content. This is in line with studies emphasizing that teacher identity is dynamic and closely tied to agency enactment, particularly in contexts of educational change. Importantly, the tourism education literature has increasingly pointed to the need for instructors to adopt expanded roles, including orchestrating immersive and experiential learning environments, often supported by technologies such as VR and AR(Ju et al., 2025; Rahmanu & Molnár, 2026). Our results provide empirical support for this identity–agency linkage in an ESL-oriented transition.

Third, and most notably, the findings show that professional identity does not directly predict pedagogical innovation (H4 not supported), whereas professional agency has a strong positive effect on pedagogical innovation (H5 supported). This distinction is theoretically and practically important. While prior studies often assume that identity change naturally leads to innovative practice, our results indicate that identity alone is insufficient; innovation requires enactment through agency. In other words, teachers may see themselves as facilitators or innovators, but without the capacity (and willingness) to act, experiment, adapt, and integrate new approaches, pedagogical innovation does not materialize. This finding refines earlier work in tourism education that highlights the benefits of experiential and technology-enhanced learning (e.g., AR/VR, simulations) (Ju et al., 2025; Takada et al., 2022) by identifying agency as the critical mechanism that turns intention into implementation.

Fourth, the strong positive effect of pedagogical innovation on communicative competence (H6 supported) confirms that experiential, simulation-based, and technology-enhanced approaches are effective in improving students' ability to communicate in authentic tourism contexts (Akyürek et al., 2024; Ju et al., 2025). This is consistent with a growing body of evidence showing that immersive learning environments, including VR-based training and interactive simulations, enhance speaking confidence, reduce anxiety, and improve real-world communication skills (Şimşek & Kalıpçı, 2023). Similarly, the use of multimedia resources and interactive tasks has been shown to support vocabulary acquisition and contextual understanding, which are essential for tourism communication. Our findings extend this literature by empirically linking pedagogical innovation, driven by teacher agency, to student-level communicative outcomes within a single integrated model.

Fifth, the explanatory and predictive metrics (R^2, Q^2, and PLS-Predict) indicate that the model has moderate to substantial explanatory power and satisfactory out-of-sample predictive capability, particularly for professional agency and pedagogical innovation (Luan-Thanh et al., 2024; Phan et al., 2025). This supports the robustness of the proposed framework and suggests that teacher-related factors are central in explaining variation in innovation and communicative competence. The effect size results further reinforce the dominance of teacher beliefs in shaping identity (large effect) and the central role of agency in driving innovation, while the negligible effect of identity on innovation corroborates the non-significant direct path.

Finally, the IPMA results provide additional managerial insight by showing that pedagogical innovation has the highest importance but the lowest performance, whereas teacher beliefs and identity demonstrate high performance but comparatively lower importance. This pattern suggests that, in the Ho Chi Minh City context, teachers already possess relatively strong communicative and industry-oriented beliefs and have begun to reconstruct their professional identities accordingly. However, the translation of these strengths into high-quality, innovative classroom practices remains insufficient. This finding resonates with earlier studies in tourism education that document persistent gaps between curricular intentions and classroom realities, often due to institutional constraints, limited resources, or lack of support for innovation (Bowen et al., 2026; Fang & Xiang, 2023). It also reinforces our central argument: agency is the bottleneck, the missing link between what teachers believe and who they think they are, and what they actually do in practice.

7. Implications

7.1 Theoretical Contributions

This study makes several important theoretical contributions that directly align with the research gaps identified in the introduction, particularly the need to understand how ESL-oriented transformation in tourism education is enacted through teacher-level mechanisms in emerging contexts such as Vietnam.

First, this study extends teacher cognition theory by situating it within the specific context of ESL-oriented tourism education, where English is not merely an academic subject but a professional communication tool. While prior research has emphasized the role of teacher beliefs in shaping instructional practices, this study advances the literature by demonstrating a sequential cognitive–behavioral mechanism in which teacher beliefs influence professional identity and professional agency, which in turn drive pedagogical innovation and ultimately affect communicative competence. This extends existing CLT and ESP research by linking cognitive orientations directly to measurable student-related outcomes in a tourism-specific ESL environment (Cloudia Ho, 2020; Ghany & Latif, 2012).

Second, the study contributes to the literature on teacher identity and professional agency by clarifying their distinct but interrelated roles in pedagogical transformation. Unlike prior studies that often treat identity as a direct predictor of innovation, this research provides empirical evidence that professional identity alone is insufficient to drive pedagogical innovation. Instead, professional agency emerges as the critical mediating mechanism that translates identity into action. This finding refines sociocultural and ecological perspectives on teacher development by highlighting that identity reconstruction must be accompanied by actionable capacity and autonomy to produce meaningful instructional change.

Third, this study enriches the growing body of research on pedagogical innovation in tourism education by integrating teacher-related constructs with emerging instructional approaches such as experiential learning, simulation-based teaching, and technology-enhanced

learning (e.g., VR/AR). While previous studies have documented the effectiveness of these innovations in improving engagement and learning outcomes (Al-Tarawneh & Osam, 2019; Rahmanu & Molnár, 2026), this study advances theory by explaining why and how such innovations are adopted, emphasizing the central role of teacher agency in enabling their implementation.

Finally, this study contributes to the broader literature on ESL transition in emerging economies, particularly Vietnam, by providing empirical evidence from a context where English is increasingly used in professional domains such as tourism. By focusing on Ho Chi Minh City, a key tourism and educational hub, the study offers a contextualized understanding of how global pedagogical shifts toward ESL are localized and enacted within specific institutional and cultural settings. This responds directly to the gap identified in the introduction regarding the lack of empirical, model-based research on ESL-oriented transformation in tourism education in developing countries.

7.2 Managerial Implications

The findings of this study offer several practical implications for policymakers, university administrators, and educators in Vietnam, particularly in aligning tourism education with industry demands and facilitating the transition toward ESL-oriented pedagogy. First, given the strong influence of teacher beliefs on identity and agency, higher education institutions should prioritize professional development programs that reshape pedagogical beliefs toward communicative, industry-relevant English. Training initiatives should move beyond traditional grammar-based instruction and focus on communicative language teaching, experiential learning, and real-world application. This is particularly important in Vietnam, where the legacy of exam-oriented education still influences teaching practices despite increasing industry demands for communication skills. Second, the results highlight that professional agency is the key driver of pedagogical innovation. Therefore, universities should create institutional environments that empower teachers to experiment with new teaching methods, redesign course content, and integrate innovative practices. This can be achieved by reducing rigid curricular constraints, encouraging pedagogical autonomy, and providing incentives for innovation. Without such support, even teachers with strong beliefs and progressive identities may be unable to implement meaningful changes in their classrooms. Third, the IPMA results indicate that pedagogical innovation has high importance but relatively low performance, suggesting a critical gap in current teaching practices. To address this, universities in Vietnam should invest in technology-enhanced learning infrastructure, such as VR/AR tools, simulation platforms, and digital learning resources. Evidence shows that such technologies can significantly enhance communicative competence and reduce learner anxiety in tourism contexts (Kusnoto et al., 2024; Rahmanu & Molnár, 2026). Additionally, incorporating industry-based simulations and experiential learning activities can help bridge the gap between classroom learning and workplace requirements. Fourth, stronger industry–university collaboration is essential. Tourism enterprises should be involved in curriculum design, guest lectures, internships, and simulation-based training to ensure that English instruction reflects real-world communication demands. This aligns with prior findings that highlight mismatches between academic training and industry expectations in tourism education (Choy & Yeung, 2024). In the Vietnamese context, particularly in Ho Chi Minh City where tourism activities are highly dynamic, such collaboration can significantly enhance the relevance and effectiveness of English language programs. Finally, policymakers at the national and institutional levels should recognize that the transition from EFL to ESL is not merely a curricular change but a systemic transformation involving teacher cognition,

institutional structures, and pedagogical practices. Strategic initiatives should therefore focus on holistic reform, including teacher training, curriculum redesign, infrastructure development, and policy support, to ensure that Vietnam's tourism education system can produce graduates with the communicative competence required in the global tourism industry.

8. Conclusion, Limitations, and Future Research

This study examined how teacher beliefs, professional identity, and professional agency jointly shape pedagogical innovation and teacher-perceived communicative competence in the transition toward ESL-oriented tourism education in Ho Chi Minh City, Vietnam. Using PLS-SEM, the analysis confirmed a sequential mechanism in which beliefs drive identity reconstruction and agency, agency—rather than identity alone—enacts pedagogical innovation, and innovation enhances teacher-perceived communicative competence. Theoretically, the study integrates teacher cognition, sociocultural, and professional agency perspectives into a single validated model and isolates agency as the pivotal link between what teachers think and what they do. Practically, the Importance–Performance Map Analysis identifies pedagogical innovation as the highest-impact yet lowest-performing lever, signaling where universities should concentrate their development efforts. In sum, advancing Vietnam's tourism English education from an EFL to an ESL orientation is less a matter of curricular documents than of empowering teachers to exercise agency in innovative practice.

Despite offering meaningful theoretical and practical insights, this study has several limitations that should be acknowledged and can guide future research. First, the study is geographically limited to Ho Chi Minh City, which, although a major educational and tourism hub in Vietnam, may not fully represent the diversity of institutional contexts across the country. Universities in other regions (e.g., Hanoi or central provinces) may differ in terms of resources, curriculum design, and industry engagement. Future research should therefore adopt a multi-region or cross-country comparative design to enhance generalizability and capture contextual variations in ESL-oriented tourism education. Second, the study relies on self-reported data from lecturers, particularly in measuring communicative competence through teacher perceptions. While this approach is common in educational research, it may introduce potential bias or overestimation of student outcomes. Future studies should incorporate multi-source data, such as student surveys, objective language performance assessments, or employer evaluations, to provide a more comprehensive and triangulated understanding of communicative competence. Third, the research adopts a cross-sectional design, which limits the ability to infer causality and capture dynamic changes over time. The transition from EFL to ESL-oriented pedagogy is inherently a longitudinal process involving gradual shifts in beliefs, identity, and practice. Future research should consider longitudinal or panel data designs to examine how teacher cognition and pedagogical innovation evolve over time, particularly in response to institutional reforms or technological adoption. Fourth, although this study focuses on key teacher-related constructs (beliefs, identity, and agency), it does not include contextual or institutional factors such as organizational support, leadership, workload, or policy constraints, which may influence teachers' ability to enact innovation. Future research could extend the model by incorporating contextual antecedents or moderating variables, such as institutional support or technological readiness, to provide a more comprehensive explanation of pedagogical transformation. Fifth, the study employs PLS-SEM to examine linear relationships, which may not fully capture complex, non-linear interactions among constructs. Given the increasing interest in advanced analytical approaches, future research could integrate Artificial Neural Networks (ANN) or hybrid SEM–ANN models to explore non-linear effects and improve predictive accuracy, particularly in understanding how

combinations of teacher characteristics influence pedagogical innovation. Finally, while this study focuses on tourism education, the proposed model may also be applicable to other ESP domains such as business English, medical English, or engineering communication. Future research could test the model in different disciplinary contexts to examine its generalizability and refine its theoretical boundaries.

References

Ab Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (2017). Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/890/1/012163

Akyürek, S., Genç, G., Çalık, İ., & Şengel, Ü. (2024). Metaverse in tourism education: A mixed method on vision, challenges and extended technology acceptance model. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *35*. https://doi.org/10.1016/j.jhlste.2024.100503

Al-Tarawneh, M. Q., & Osam, U. V. (2019). Tourism English training at the tertiary level in Jordan: Reality and expectations from a university context. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *24*, 155–167. https://doi.org/10.1016/j.jhlste.2019.02.005

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. Language Teaching, 36(2), 81–109. https://doi.org/10.1017/S0261444803001903

Bowen, Z., Runpu, Y., Yuemin, T., & Shengfeng, Z. (2026). Elements of tourist demand for tourism-related performing arts based on online reviews and the Kano model. *Entertainment Computing*, *56*, 101081. https://doi.org/10.1016/j.entcom.2026.101081

Choy, M. W., & Yeung, A. S. (2024). Tourism vocational education: Relations of input and output. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *34*. https://doi.org/10.1016/j.jhlste.2024.100479

Cloudia Ho, Y. Y. (2020). Communicative language teaching and English as a foreign language undergraduates' communicative competence in Tourism English. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *27*. https://doi.org/10.1016/j.jhlste.2020.100271

Elfiondri, Zaitul, & Rina, N. (2021). Tradition, cultural contact and English for tourism: the case of Mentawai, Indonesia. *Heliyon*, *7*(6). https://doi.org/10.1016/j.heliyon.2021.e07322

Fang, X., & Xiang, K. (2023). Tourism visual deception in "check-in tourism": Value co-creation and co-destruction perspectives. *Journal of Hospitality and Tourism Management*, *57*, 190–199. https://doi.org/10.1016/j.jhtm.2023.10.009

Ghany, S. Y. A., & Latif, M. M. A. (2012). English language preparation of tourism and hospitality undergraduates in Egypt: Does it meet their future workplace requirements? *Journal of Hospitality, Leisure, Sport and Tourism Education*, *11*(2), 93–100. https://doi.org/10.1016/j.jhlste.2012.05.001

Giovanis, A., Rizomyliotis, I., Konstantoulaki, K., & Magrizos, S. (2021). Mining the hidden seam of proximity m-payment adoption: A hybrid PLS-artificial neural network analytical approach. *European Management Journal*, (March). https://doi.org/10.1016/j.emj.2021.09.007

Habib, J., Prentice, C., & Thaichon, P. (2026). Employee–robot relationship typology: the role of robot agency, communion and competence. *Tourism Review*, 1–22. https://doi.org/10.1108/TR-08-2025-0872

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152.

He, Z., Hu, Z. A., Huang, W., & Kang, Y. (2025). Tourism growth, education decline: Evidence from China's 5A attraction expansion. *Journal of Urban Economics*, *150*. https://doi.org/10.1016/j.jue.2025.103811

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-014-0403-8

Joseph F. Hair, J., Marko, S., Christian M. Ringle, & Siegfried P. Gudergan. (2023). *Advanced Issues in Partial Least Squares Structural Equation Modeling*. SAGE Publications. https://uk.sagepub.com/en-gb/eur/advanced-issues-in-partial-least-squares-structural-equation-modeling/book279526

Ju, Y., Tsai, Y. Y., Gao, S. Y., Wu, T. liang, & Tseng, T. L. (2025). Integrating augmented reality into event tourism education: Enhancing experiential value and authenticity. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *37*. https://doi.org/10.1016/j.jhlste.2025.100571

Kacetl, J., & Frydrychova-Klimova, B. (2015). English Vocabulary in Video Clips on Travel and Tourism. *Procedia - Social and Behavioral Sciences*, *182*, 364–368. https://doi.org/10.1016/j.sbspro.2015.04.788

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Kusnoto, Y., Supriatna, N., Wiyanarti, E., & Hasan, S. H. (2024). Trend and visualizing of historical tourism in education research during last twenty years: A bibliometric review and analysis. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *34*. https://doi.org/10.1016/j.jhlste.2023.100477

Langum, V., & Sullivan, K. P. H. (2020). Academic writing, scholarly identity, voice and the benefits and challenges of multilingualism: Reflections from Norwegian doctoral researchers in teacher education. *Linguistics and Education*. https://doi.org/10.1016/j.linged.2020.100883

Liu, C. H., Horng, J. S., Chou, S. F., Yu, T. Y., Huang, Y. C., Ng, Y. L., & La, Q. P. (2024). Explore links among marketing knowledge, data literacy, skill improvement, and learning attitude in STEAM application for hospitality and tourism education. *International Journal of Management Education*, *22*(1). https://doi.org/10.1016/j.ijme.2023.100919

Luan-Thanh, N., Dang, T.-Q., & Nguyen, T.-T. T. (2024). *Metaverse in Education: Bridging Realities for Future Learning* (N. Luan-Thanh, Ed.; Vol. 1). Generis Publishing. https://www.generis-publishing.com/book.php?title=strong-metaverse-in-education-bridging-realities-for-future-learning-strong-2499

Ooi, K. B., & Tan, G. W. H. (2016). Mobile technology acceptance model: An investigation using mobile users to explore smartphone credit card. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2016.04.015

Pantaruk, S., Khuadthong, B., Sangthong, M., Srirat, P., Imjai, N., Suhud, U., Boonkaew, S., & Aujirapongpan, S. (2026). Transforming historical literacy into sustainable cultural tourism behavior: The mediating role of social and critical thinking in tourism education. *Sustainable Futures*. https://doi.org/10.1016/j.sftr.2026.101794

Phan, T. – T. C., Nguyen, L.-T., & Dang, T.-Q. (2025). Does Impulsive Buying Lead to Compulsive Buying in Metaverse? A Dual-Stage Predictive-Analytics SEM-ANN Analysis. The Empirical Study from Vietnam. *Thailand and The World Economy*, *43*(2), 44–46. https://so05.tci-thaijo.org/index.php/TER/article/view/268273

Priestley, M., Biesta, G., & Robinson, S. (2015). Teacher agency: An ecological approach. Bloomsbury Academic.

Rahmanu, I. W. E. D., & Molnár, G. (2026). Virtual reality in English language education: A systematic review of opportunities and challenges in higher education. In *Acta Psychologica* (Vol. 264). Elsevier B.V. https://doi.org/10.1016/j.actpsy.2026.106358

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. *European Journal of Marketing*. https://doi.org/10.1108/EJM-02-2019-0189

Şimşek, E. K., & Kalıpçı, M. B. (2023). A bibliometric study on higher tourism education and curriculum. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *33*. https://doi.org/10.1016/j.jhlste.2023.100442

Takada, K., Murakami, K. H., Yamazaki, A. K., & Yamanaka, T. (2022). A preliminary study of VR English training material for personnel in the tourism industry. *Procedia Computer Science*, *207*, 3715–3723. https://doi.org/10.1016/j.procs.2022.09.432

Tao, J., & Gao, X. (2017). Teacher agency and identity commitment in curricular reform. *Teaching and Teacher Education*, *63*, 346–355. https://doi.org/10.1016/j.tate.2017.01.010

Wyatt, M., & Dikilitaş, K. (2019). English language teachers' self-efficacy beliefs for grammar instruction: implications for teacher educators. *The Language Learning Journal*. https://doi.org/10.1080/09571736.2019.1642943

#196 FROM TRADITIONAL TESTING TO ALTERNATIVE ASSESSMENT: EXAMINING THE RELATIONSHIPS AMONG CLASSROOM-BASED ALTERNATIVE ASSESSMENT, STUDENT ENGAGEMENT, AND PERCEIVED COMMUNICATIVE COMPETENCE IN A VIETNAMESE ESL HIGHER-EDUCATION CONTEXT

Phan Tran Duc Lien[1]

Abstract:

The transition toward communicative and learner-centered English language education has intensified the need to move beyond traditional examination-oriented assessment practices toward more dynamic classroom-based alternative assessment approaches. Despite growing recognition of formative and authentic assessment, empirical evidence regarding the relationships among alternative assessment, student engagement, and students' perceived communicative competence remains limited, particularly in emerging educational contexts such as Vietnam. Drawing upon constructivist learning perspectives and engagement theory, this study develops and tests a research model examining these relationships. Survey data were collected from 412 undergraduate students enrolled in English-related courses in Ho Chi Minh City, Vietnam, and the proposed model was analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM). The findings indicate that classroom-based alternative assessment is positively associated with student engagement and perceived communicative competence. Student engagement is also positively associated with perceived communicative competence, highlighting the importance of active participation and interaction in language learning environments. The model demonstrates satisfactory explanatory power and cross-validated predictive relevance. However, because the study employs a cross-sectional survey design and self-reported measures, the findings should be interpreted as predictive associations rather than definitive causal effects or objective evidence of linguistic performance. This study contributes to the literature by conceptualizing classroom-based alternative assessment as a learning-enabling mechanism and integrating assessment practices, engagement, and perceived communicative outcomes within a unified framework. The practical implications outline how educators and institutions can embed alternative assessment into existing learning management systems and departmental assessment practices.

Keywords: Classroom-based alternative assessment, Student engagement, Perceived communicative competence, English language education, Language learning assessment, Higher education

[1] *Faculty of Business Administration, Ho Chi Minh City University of Foreign Languages-Information Technology*

Email: lienptd@huflit.edu.vn

1. Introduction

The rapid globalization of education and labor markets has elevated English proficiency, particularly communicative competence, as a critical capability for individuals operating in multilingual and multicultural environments (Unesco, 2025). In this context, English as a Second Language (ESL) education has undergone a paradigm shift from grammar-based instruction toward communicative, interaction-driven learning models (Fonteles et al., 2026; Guilloteaux, 2013). However, despite these pedagogical advancements, assessment practices in many ESL systems continue to rely heavily on traditional, examination-oriented approaches that emphasize memorization and static knowledge reproduction rather than authentic language use and communicative performance.

This misalignment between instructional goals and assessment practices represents a fundamental challenge in contemporary education. Traditional assessment methods are increasingly criticized for their inability to capture the dynamic, socially constructed nature of language learning, particularly in contexts where communication, interaction, and real-world application are essential (Fink et al., 2026). In contrast, classroom-based assessment (CLB) and alternative assessment approaches, such as formative assessment, peer evaluation, and portfolio-based assessment, have been proposed as more effective mechanisms for integrating assessment into the learning process (Badem, 2025; Muhonen et al., 2025). These approaches emphasize continuous feedback, learner participation, and contextualized performance, thereby aligning assessment with constructivist learning principles.

Empirical evidence supports the effectiveness of such approaches. For instance, formative and classroom-integrated assessment practices have been shown to enhance student engagement and promote deeper learning experiences beyond traditional examination formats (Sanford & Snow, 2026). Similarly, process-oriented classroom evaluation, including behavioral observation and real-time feedback, provides more comprehensive insights into student learning and instructional effectiveness (Yu et al., 2026). In addition, experiential and interactive learning environments have been found to significantly improve higher-order competencies, including communication, creativity, and analytical skills (Chopra, 2025). These findings collectively suggest that alternative assessment approaches are not merely evaluative tools but transformative pedagogical mechanisms that shape learning processes and outcomes.

A critical mechanism through which alternative assessment influences learning outcomes is student engagement. Engagement is widely conceptualized as a multidimensional construct encompassing behavioral participation, emotional involvement, and cognitive investment in learning activities (Mele et al., 2025). It plays a central role in determining academic success, as engaged students are more likely to actively participate, persist in learning tasks, and achieve higher performance outcomes. Conversely, limited engagement has been consistently associated with reduced academic achievement and diminished learning effectiveness, particularly in technology-mediated and virtual learning environments (Rabby et al., 2025). Recent research further highlights that engagement is shaped by interactive, feedback-rich, and student-centered learning environments, reinforcing the importance of integrating assessment practices that actively involve learners.

At the same time, the ultimate goal of English language education, communicative competence, has evolved into a multidimensional construct encompassing linguistic proficiency, intercultural awareness, adaptability, and strategic communication skills. Contemporary frameworks conceptualize perceived communicative competence as a combination of knowledge, attitudes, and skills that enable individuals to interact effectively in diverse contexts (Ertay & Gilanlioglu, 2024). In an increasingly globalized world,

intercultural communication competence has become a fundamental requirement for academic, professional, and social success, with deficiencies in such competence often leading to failure in cross-cultural interactions and professional environments (Sarwari et al., 2024). Moreover, communication competence extends beyond language proficiency to include empathy, adaptability, and the ability to tailor communication to different audiences and contexts (Fuller et al., 2021), as well as the ability to operate effectively in technology-mediated environments (Koponen et al., 2025).

Despite the recognized importance of alternative assessment, student engagement, and perceived communicative competence, existing research remains fragmented. Prior studies have typically examined these constructs in isolation, focusing either on assessment practices, engagement dynamics, or communication outcomes, without integrating them into a comprehensive theoretical model (Ahmed et al., 2026; Rahmanu & Molnár, 2026; Sieveke et al., 2026; Xu et al., 2026). Furthermore, there is a notable lack of empirical research applying advanced analytical techniques such as Partial Least Squares Structural Equation Modeling (PLS-SEM) to examine these relationships in ESL contexts, particularly in emerging economies such as Vietnam where educational transformation is ongoing. In addition, contemporary educational trends, including digital transformation, AI integration, and active learning methodologies, further underscore the need to rethink assessment practices. Technologies such as AI-driven learning systems and interactive tools are reshaping student engagement and learning experiences (Mele et al., 2025), while active learning approaches and gamified assessment have been shown to significantly enhance engagement and academic performance compared to traditional examination methods (Roig et al., 2026). Similarly, alternative grading approaches, such as upgrading and learner-centered evaluation, have been found to reduce anxiety and shift students' focus toward meaningful learning rather than performance metrics (Burns et al., 2026). These developments highlight the urgency of developing integrated frameworks that capture the complex interplay between assessment, engagement, and learning outcomes in modern educational contexts.

To address these gaps, the present study develops and empirically tests a comprehensive research model examining the interconnected roles of classroom-based alternative assessment, student engagement, and perceived communicative competence. By integrating insights from constructivist learning theory, engagement research, and communication competence frameworks, this study contributes to the literature in three key ways. First, it advances theoretical understanding by linking assessment practices, student engagement, and communicative outcomes within an integrated framework. Second, it provides empirical evidence from an emerging ESL context using advanced analytical methods (PLS-SEM). Third, it offers practical implications for educators and policymakers seeking to reform assessment practices to better align with the demands of 21st-century language education.

2. Literature Review

2.1. Classroom-Based and Alternative Assessment in ESL Contexts

The shift from traditional, examination-driven assessment toward classroom-based and alternative assessment (CLB) represents a fundamental transformation in contemporary education, particularly in ESL contexts where communicative competence is a primary learning objective. Traditional assessment approaches, characterized by standardized testing and summative evaluation (Burns et al., 2026; Lee & Jeon, 2024; Sanford & Snow, 2026), have been widely criticized for their limited ability to capture the dynamic, contextual, and interactional nature of language learning. These approaches often prioritize memorization and discrete linguistic knowledge, thereby neglecting higher-order competencies such as

communication, collaboration, and real-world language use. In response to these limitations, classroom-based assessment has emerged as a learner-centered paradigm that integrates assessment into the teaching–learning process (Lee & Jeon, 2024). CLB emphasizes continuous feedback, contextualized performance, and authentic learning tasks, enabling teachers to monitor student progress and adapt instruction accordingly (Sanford & Snow, 2026). Within this paradigm, alternative assessment methods, including formative assessment, peer assessment, and portfolio-based assessment, play a central role in fostering active learning and meaningful engagement.

Although technology-mediated assessment is increasingly important in contemporary ESL education, the present study focuses on the pedagogical foundations of classroom-based alternative assessment rather than on a specific technological platform. The CLB construct captures formative, interactive, process-oriented, and authentic assessment practices, including continuous feedback, peer evaluation, portfolios, and classroom discussions. These practices can be implemented through conventional classroom activities or supported by digital platforms such as learning management systems, e-portfolios, online discussion forums, automated quizzes, and AI-assisted feedback tools. Future research should extend the present framework by explicitly measuring technology-mediated assessment practices and examining whether digital readiness or AI-supported feedback strengthens the relationships among alternative assessment, engagement, and perceived communicative competence.

Empirical evidence consistently demonstrates the pedagogical value of alternative assessment approaches. For instance, formative assessment practices that incorporate continuous feedback and reinforcing activities have been shown to enhance student engagement and deepen learning beyond traditional examination formats (Sanford & Snow, 2026). Similarly, process-oriented evaluation systems enable teachers to capture real-time behavioral indicators of learning, thereby improving both instructional quality and student outcomes (Yu et al., 2026). These findings align with broader research on experiential learning, which highlights the importance of integrating real-world tasks and authentic experiences into educational environments to enhance higher-order competencies such as communication, creativity, and problem-solving (Chopra, 2025).

Furthermore, recent developments in educational technology and digital pedagogy have reinforced the importance of alternative assessment. The integration of interactive tools, game-based systems, and AI-driven learning environments has enabled more immediate feedback, adaptive learning, and inclusive participation, all of which contribute to improved learning experiences and outcomes (Priante & Tsekouras, 2025). At the same time, innovative assessment formats, such as gamified evaluation and escape-room-based assessment, have demonstrated significant improvements in student performance and engagement compared to traditional testing methods (Chopra, 2025). These advancements underscore the evolving role of assessment as a dynamic and integral component of the learning process rather than a static measure of performance. Despite these benefits, the adoption of classroom-based alternative assessment remains uneven, particularly in emerging ESL contexts where institutional structures and cultural norms continue to favor examination-oriented systems. This tension highlights the need for empirical research that examines how alternative assessment practices operate within specific educational contexts and how they influence key learning outcomes such as engagement and perceived communicative competence.

2.2. Student Engagement in Contemporary Learning Environments

Student engagement has become a central construct in educational research, reflecting its critical role in shaping learning processes and outcomes (Byers et al., 2024). Engagement is

commonly conceptualized as a multidimensional construct encompassing behavioral participation, emotional involvement, and cognitive investment in learning activities (Carey, 2013). This multidimensionality highlights that engagement is not merely about participation but also about the depth of learners' involvement and their commitment to learning.

The importance of engagement is particularly evident in language learning contexts, where active participation and interaction are essential for developing perceived communicative competence. Research indicates that engagement is closely linked to students' motivation, persistence, and academic performance. For example, studies have shown that limited engagement in learning environments, especially in technology-mediated contexts, can significantly reduce academic performance and learning effectiveness (Rabby et al., 2025). Conversely, interactive and feedback-rich learning environments have been found to enhance students' cognitive and emotional involvement, leading to improved learning outcomes.

Recent research further elaborates the mechanisms through which engagement operates in contemporary educational settings. Advances in digital learning technologies, including AI-driven tools and educational robots, have demonstrated the potential to enhance engagement by providing personalized, interactive, and multisensory learning experiences. These technologies can simultaneously influence cognitive, emotional, and behavioral dimensions of engagement, thereby creating more holistic and effective learning environments (Sahlender et al., 2024). Similarly, human-centered approaches to educational innovation emphasize the importance of student agency, participation, and co-creation in shaping meaningful learning experiences (O'Brien et al., 2026).

Moreover, engagement is not only an individual-level phenomenon but also a socially constructed process influenced by classroom interactions, instructional practices, and institutional contexts. Research on classroom interactions highlights the importance of communicative participation, collaborative learning, and teacher–student interaction in fostering engagement and supporting language acquisition (Badem, 2025; Kozaklı Ülger et al., 2026). Additionally, co-creation and participatory learning approaches have been shown to enhance students' sense of belonging, motivation, and commitment to learning, thereby reinforcing engagement as a key driver of educational success (O'Brien et al., 2026).

Within this broader context, classroom-based alternative assessment can be understood as a critical mechanism for shaping student engagement. By providing continuous feedback, promoting interaction, and emphasizing authentic tasks, alternative assessment practices create conditions that encourage active participation and deeper involvement in learning. This perspective suggests that engagement is not merely an outcome but an integral component of the assessment–learning nexus.

2.3. Perceived communicative competence as a Core Outcome in ESL Education

Communicative competence has long been recognized as a central objective of English language education, reflecting learners' ability to use language effectively in real-world contexts (Cloudia Ho, 2020). Unlike traditional measures of language proficiency that focus primarily on grammatical accuracy and vocabulary knowledge, communicative competence encompasses a broader set of capabilities, including sociolinguistic appropriateness, discourse management, strategic communication, and intercultural awareness (Ertay & Gilanlioglu, 2024).

The present study focuses specifically on perceived communicative competence, which refers to students' self-evaluated confidence and perceived ability to communicate effectively in English. This distinction is important because students' perceptions may influence their

willingness to participate, interact, and engage in communicative activities, but they do not necessarily reflect objectively demonstrated linguistic performance. Accordingly, the empirical construct examined in this study should be interpreted as perceived communicative competence rather than as an objective measure of English proficiency.

2.4. Linking Alternative Assessment, Engagement, and Perceived communicative competence

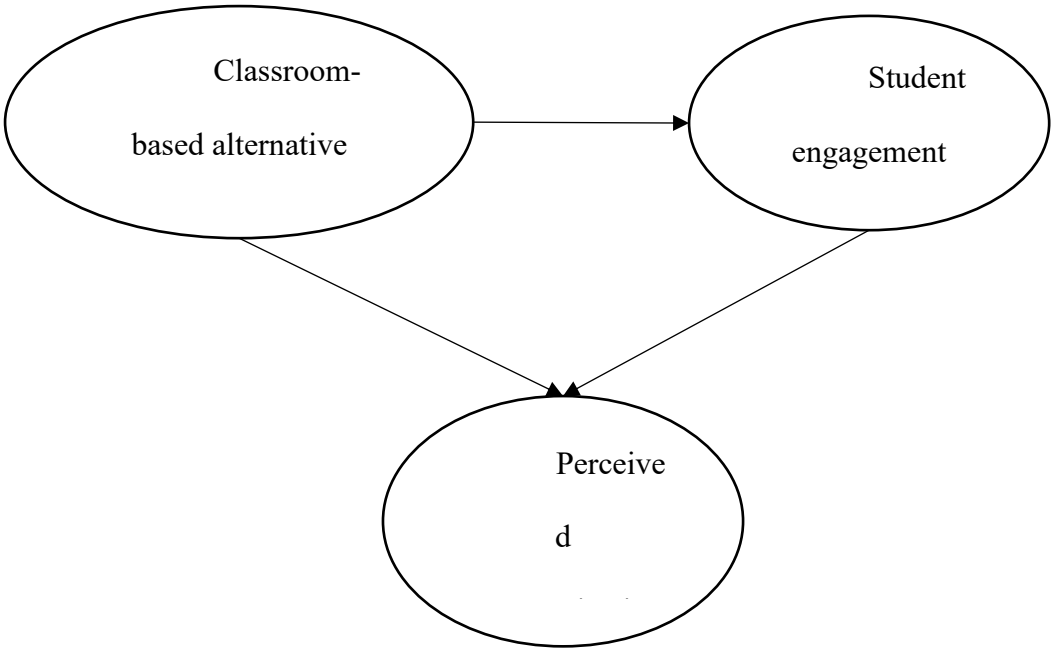
The preceding discussions highlight that classroom-based alternative assessment, student engagement, and perceived communicative competence are deeply interconnected within the ESL learning process. Alternative assessment practices provide the structural and pedagogical conditions for active learning, engagement reflects the degree of learner involvement and investment, and perceived communicative competence represents the ultimate outcome of language education. Rather than treating these constructs as isolated components, contemporary educational research increasingly emphasizes their integration within a holistic learning framework. Alternative assessment practices, by emphasizing feedback, interaction, and authenticity, create learning environments that encourage engagement and support the development of perceived communicative competence. Engagement, in turn, reflects the extent to which learners actively participate in and benefit from these environments, while perceived communicative competence captures the effectiveness of learning in achieving real-world communication outcomes. Despite the theoretical alignment among these constructs, empirical research integrating them into a unified framework remains limited, particularly in ESL contexts within emerging economies. This gap is further compounded by the lack of studies employing advanced analytical techniques such as PLS-SEM to examine the relationships among assessment practices, engagement, and communication outcomes. To address these gaps, the present study focuses specifically on perceived communicative competence, which refers to students' self-evaluated confidence and perceived ability to communicate effectively in English. This distinction is important because students' perceptions may influence their willingness to participate, interact, and engage in communicative activities, but they do not necessarily represent objectively demonstrated linguistic performance. Accordingly, the empirical construct examined in this study should be interpreted as perceived communicative competence rather than as an objective measure of English proficiency. Based on the literature reviewed, the following hypotheses are proposed:

H1: Classroom-based alternative assessment is positively associated with student engagement.

H2: Classroom-based alternative assessment is positively associated with perceived communicative competence.

H3: Student engagement is positively associated with perceived communicative competence.

The research model is presented in Figure 1.



3. Methodology

3.1 Research Design

This study adopts a quantitative research design to examine the relationships among classroom-based alternative assessment, student engagement, and perceived communicative competence in ESL contexts. A cross-sectional survey method was employed, as it is widely used in educational and behavioral research to capture students' perceptions and experiences in natural learning environments. Given the complexity of the proposed model and the predictive orientation of the study, Partial Least Squares Structural Equation Modeling (PLS-SEM) was selected as the primary analytical technique. PLS-SEM is particularly suitable for exploratory and prediction-oriented research, as well as for models involving multiple latent constructs and complex relationships.

3.2 Sample and Minimum Sample Size

The target population of this study consists of undergraduate students enrolled in English-related courses at universities in Ho Chi Minh City, Vietnam. These students represent an appropriate sample because they are directly exposed to classroom-based assessment practices and are actively engaged in ESL learning environments. Following established guidelines for PLS-SEM, the minimum sample size was determined using the "10-times rule", which suggests that the sample size should be at least ten times the maximum number of structural paths pointing at any construct in the model. In the present study, the maximum number of arrows pointing to a construct is two (from CLB and STE to PCC), resulting in a minimum sample size of 20. However, this rule is considered conservative. To ensure statistical power and robustness, this study follows more rigorous recommendations suggesting a minimum sample size of 200 for SEM-based analysis. Accordingly, a total of 450 questionnaires were distributed, and 412 valid responses were retained after data screening, exceeding recommended thresholds and ensuring sufficient statistical power for PLS-SEM analysis.

3.3 Data Collection Procedure

Data were collected using a structured questionnaire administered to students in ESL courses. The survey was distributed both online and offline to increase response rates and ensure diverse participation. Prior to the main data collection, a pilot test with 30 students was conducted to assess the clarity, reliability, and validity of the measurement items. Participation was voluntary, and respondents were assured of anonymity and confidentiality. Screening questions were used to ensure that participants had prior experience with English courses and classroom-based assessment practices. All constructs were measured using a 7-point Likert scale ranging from: (1) Strongly disagree → (7) Strongly agree. The use of a 7-point scale is recommended in SEM research as it provides greater variance, improves measurement sensitivity, and enhances reliability. All measurement items were adapted and refined from prior validated studies on classroom-based assessment, student engagement, and perceived communicative competence, with modifications to fit the ESL context.

3.4 Measurement Items

Table 1

Measurement items

Constructs	**Definition**	**Items**	**References**
Classroom-Based Alternative Assessment (CLB)	CLB reflects formative, interactive, and authentic assessment practices aligned with classroom-based assessment literature	CLB1. Assessment activities are integrated into classroom learning tasks.	Sanford & Snow (2026)
		CLB2. I receive continuous feedback during English learning activities.	
		CLB3. My teacher uses various assessment methods (e.g., peer, portfolio, discussion).	
		CLB4. Assessment focuses on my learning process rather than only final results.	
		CLB5. Classroom assessment helps me understand my strengths and weaknesses.	
Student Engagement (STE)	STE reflects behavioral, emotional, and cognitive engagement dimensions	STE1. I actively participate in English classroom activities.	Mele et al. (2025)
		STE2. I feel interested in learning English during class.	
		STE3. I am motivated to complete English learning tasks.	
		STE4. I pay attention during English lessons.	

		STE5. I enjoy interacting with classmates in English.	
Perceived communicative competence (PCC)	Perceived communicative competence (PCC) refers to students' self-evaluated confidence and perceived ability to express ideas, participate in discussions, respond appropriately in conversations, and use English in real-life situations. The construct captures learners' subjective evaluations of their communicative capability rather than objectively assessed linguistic performance.	PCC1. I can express my ideas clearly in English.	Ertay & Gilanlioglu (2024)
		PCC2. I feel confident communicating in English.	
		PCC3. I can participate in discussions using English.	
		PCC4. I can understand and respond appropriately in conversations.	
		PCC5. I can use English effectively in real-life situations.	

The present CLB scale captures the pedagogical foundations of classroom-based alternative assessment and does not directly operationalize technology-mediated or AI-supported assessment practices.

3.5 Common Method Bias (CMB)

Given that all data were collected from a single source using self-reported measures, potential common method bias (CMB) was addressed through both procedural and statistical remedies.

Procedural Remedies (Kock, 2015)

- Anonymity and confidentiality were assured to reduce evaluation apprehension
- Items were carefully worded to avoid ambiguity
- Different constructs were separated conceptually in the questionnaire

Statistical Remedies

First, Harman's single-factor test was conducted, and the first factor accounted for less than 50% of the total variance, indicating that CMB is unlikely to be a serious concern (Harman, 1976). Second, full collinearity assessment (VIF) was performed in Table 3, with all VIF values below the threshold of 3.3, further confirming the absence of common method bias (Thi Viet & Nguyen, 2025).

4. Results, Discussion, Implications

4.1 Results

4.1.1 Sample Characteristics

A total of 412 valid responses were retained for analysis in Table 2. The demographic profile indicates a relatively balanced distribution across gender, academic levels, and majors,

ensuring adequate representation of students with different levels of exposure to ESL learning environments. Female respondents accounted for the majority (56.3%), followed by males (40.8%), while a small proportion identified as other (2.9%). In terms of academic level, most respondents were in their second (30.1%) and third years (33.0%), suggesting sufficient experience with classroom-based assessment practices. Regarding academic majors, Business Administration (37.9%) and English Studies (34.5%) represented the largest groups. Additionally, a substantial proportion of students reported frequent use of English in class, with 64.1% indicating "often" or "very often," reflecting a communicative learning environment suitable for this study.

Table 2

Demographic Profile of Respondents (N = 412)

Category	Item	Frequency (n)	Percentage (%)
Gender	Male	168	40.8
	Female	232	56.3
	Other	12	2.9
Year of Study	1st Year	80	19.4
	2nd Year	124	30.1
	3rd Year	136	33.0
	4th Year	72	17.5
Major	Business Administration	156	37.9
	English Studies	142	34.5
	Others	114	27.6
Frequency of English Use in Class	Rarely	36	8.7
	Sometimes	112	27.2
	Often	168	40.8
	Very often	96	23.3

4.1.2 Measurement model assessment

Table 3 presents the results of convergent validity and internal consistency assessment for all constructs, including classroom-based alternative assessment (CLB), student engagement (STE), and perceived communicative competence (PCC). Overall, the results demonstrate that all constructs meet the recommended thresholds for reliability and validity. First, the indicator loadings range from 0.704 to 0.868, exceeding the minimum acceptable threshold of 0.70 (Hair et al., 2019), thereby confirming satisfactory indicator reliability. Although CLB5 exhibits a relatively lower loading (0.704), it remains acceptable and contributes to content validity. In terms of internal consistency, Cronbach's alpha values range from 0.838 (CLB) to 0.887 (PCC), exceeding the recommended threshold of 0.70, indicating strong reliability. Similarly, composite reliability values (rho_a and rho_c) for all constructs exceed 0.80, further confirming the robustness of the measurement model. Specifically, rho_c values range from 0.885 to 0.917, demonstrating high internal consistency across constructs. Convergent validity is assessed using the Average Variance Extracted (AVE), with all

constructs exceeding the threshold of 0.50 (Ferrell et al., 2019). The AVE values are 0.607 for CLB, 0.686 for STE, and 0.690 for PCC, indicating that each construct explains more than half of the variance of its indicators. This confirms that the measurement items adequately represent their respective latent constructs.

Table 3

Convergent validity, Internal consistency assessment

Constructs	Items	Loadings	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)	VIF
CLB	**CLB1**	0.794	0.838	0.840	0.885	0.607	1.919
	CLB2	0.771					1.804
	CLB3	0.796					1.926
	CLB4	0.825					2.120
	CLB5	0.704					1.304
PCC	**PCC1**	0.861	0.887	0.889	0.917	0.690	2.620
	PCC2	0.829					2.283
	PCC3	0.865					2.524
	PCC4	0.829					2.191
	PCC5	0.766					1.739
STE	**STE1**	0.776	0.885	0.886	0.916	0.686	1.758
	STE2	0.828					2.131
	STE3	0.845					2.712
	STE4	0.868					2.827
	STE5	0.821					2.109

Discriminant validity was assessed using the Heterotrait–Monotrait ratio (HTMT), as presented in Table 4. The HTMT values among all constructs are below the conservative threshold of 0.80 (Ab Hamid et al., 2017), indicating satisfactory discriminant validity. Specifically, the HTMT value between CLB and PCC is 0.648, between CLB and STE is 0.767, and between STE and PCC is 0.766. All values are well below the threshold, suggesting that each construct is empirically distinct and captures unique aspects of the conceptual model. These findings confirm that classroom-based alternative assessment, student engagement, and perceived communicative competence are not only theoretically distinct but also empirically separable constructs. Therefore, the measurement model satisfies the requirements for discriminant validity.

Table 4

Discriminant validity (HTMT 0.80)

	CLB	PCC	STE

CLB			
PCC	0.648		
STE	0.767	0.766	

4.1.3 Structural model assessment

Table 5 presents the results of the structural model, including path coefficients, t-values, p-values, and confidence intervals obtained from the bootstrapping procedure (5,000 resamples). Overall, the structural model demonstrates strong explanatory power, with all hypothesized relationships being positive and statistically significant. The results provide robust empirical support for the proposed relationships among classroom-based alternative assessment (CLB), student engagement (STE), and perceived communicative competence (PCC). First, the effect of classroom-based alternative assessment on student engagement is positive and highly significant ($\beta = 0.677$, $t = 19.988$, $p < 0.001$). The confidence interval [0.607; 0.741] does not include zero, further confirming the robustness of the relationship. This result indicates that classroom-based alternative assessment is positively associated with students' active participation, motivation, and involvement in ESL learning activities. Among all relationships, this path exhibits the strongest effect size, highlighting the central role of assessment practices in shaping engagement. Second, classroom-based alternative assessment also has a direct and statistically significant effect on perceived communicative competence ($\beta = 0.201$, $t = 3.872$, $p < 0.001$). Although the magnitude of this relationship is smaller compared to other paths, the confidence interval [0.098; 0.300] confirms its significance. This finding suggests that alternative assessment practices are positively associated with perceived communicative competence by providing authentic, feedback-rich learning experiences that reflect real-world language use. Third, student engagement has a strong and significant effect on perceived communicative competence ($\beta = 0.542$, $t = 11.057$, $p < 0.001$), with a confidence interval of [0.444; 0.637]. This result demonstrates that higher levels of engagement, manifested through active participation, motivation, and interaction, are associated with improved communicative abilities. The magnitude of this relationship indicates that engagement is an important predictor of perceived communicative competence in ESL contexts. Taken together, the findings confirm that classroom-based alternative assessment not only is positively associated with perceived communicative competence but also is strongly associated with student engagement, which in turn is closely associated with communicative outcomes. All hypotheses (H1, H2, and H3) are therefore supported.

Table 5

Structural model results

Hypothesis	**Path**	**Original sample (O)**	**Sample mean (M)**	**Standard deviation (STDEV)**	**T statistics (|O/STDEV|)**	**P values**	**2.50 %**	**97.50 %**	**Remarks**
H1	CLB -> STE	0.677	0.677	0.034	19.988	0.000	0.607	0.741	Supported

H2	CLB -> PCC	0.201	0.201	0.052	3.872	0.000	0.098	0.300	Supported
H3	STE -> PCC	0.542	0.543	0.049	11.057	0.000	0.444	0.637	Supported

Table 6 presents the coefficient of determination (R^2) and predictive relevance (Q^2) for the endogenous constructs, namely perceived communicative competence (PCC) and student engagement (STE). The R^2 values indicate the explanatory power of the structural model. Specifically, the R^2 value for PCC is 0.482, suggesting that 48.2% of the variance in PCC is explained by classroom-based alternative assessment (CLB) and student engagement (STE). According to commonly accepted benchmarks, this value can be interpreted as moderate explanatory power, indicating that the model provides substantial insight into the determinants of PCC in ESL contexts. Similarly, the R^2 value for student engagement (STE) is 0.458, meaning that 45.8% of the variance in engagement is explained by classroom-based alternative assessment. This result also reflects moderate explanatory capability, highlighting the strong role of assessment practices in shaping student engagement within classroom environments. In addition to explanatory power, the predictive relevance of the model was assessed using the Stone–Geisser Q^2 value obtained through the blindfolding procedure. The Q^2 value for PCC (PCC) is 0.328, while the Q^2 value for student engagement (STE) is 0.305. Both values are well above zero, indicating that the model has strong predictive relevance for both endogenous constructs (Stone, 1974). More importantly, Q^2 values above 0.25 are generally considered to indicate medium to high predictive accuracy, suggesting that the model is not only explanatory but also capable of predicting out-of-sample observations effectively. This reinforces the robustness and practical applicability of the proposed model in ESL educational settings.

Table 6

R-square and Q square

	R-square	**SSO**	**SSE**	**Q^2 (=1-SSE/SSO)**
PCC	0.482	2060	1383.481	0.328
STE	0.458	2060	1431.379	0.305

4.2 Discussion

This study set out to examine how CLB, STE, and PCC are interconnected in ESL learning environments. The findings provide strong empirical support for all proposed relationships and offer several important insights into the role of assessment practices in shaping learning outcomes.

4.2.1 The Association between Classroom-Based Alternative Assessment and Student Engagement

The results reveal that classroom-based alternative assessment has a strong and significant effect on student engagement ($\beta = 0.677$), indicating that assessment practices

embedded within classroom activities substantially enhance students' active participation, motivation, and involvement. This finding underscores the central role of assessment as a pedagogical tool rather than merely an evaluative mechanism. This result is consistent with prior research suggesting that formative and process-oriented assessment approaches foster deeper engagement by providing continuous feedback and encouraging active learning behaviors (Sanford & Snow, 2026). Similarly, classroom-based evaluation systems that emphasize real-time interaction and behavioral observation have been shown to promote greater student involvement and responsiveness in learning environments (Yu et al., 2026). The present findings extend this line of research by providing quantitative evidence, using PLS-SEM, that classroom-based alternative assessment is a key driver of engagement in ESL contexts. Moreover, the magnitude of this relationship aligns with emerging studies on interactive and technology-enhanced learning environments, which highlight that engagement is strongly influenced by feedback-rich, participatory, and student-centered instructional practices (Mele et al., 2025). This suggests that alternative assessment practices function as an important mechanism through which such environments are operationalized in classroom settings.

4.2.2 The Association between Alternative Assessment and Perceived Communicative Competence

The findings also demonstrate that classroom-based alternative assessment has a direct and significant effect on perceived communicative competence ($\beta = 0.201$). Although this effect is smaller than its influence on engagement, it remains statistically meaningful and theoretically important. This result supports the argument that assessment practices grounded in authentic, real-world tasks contribute directly to the development of communicative abilities. Previous studies have emphasized that experiential and interactive learning environments enhance communication-related competencies by allowing students to apply language in practical contexts (Sanford & Snow, 2026). Similarly, research on active learning and innovative assessment methods, such as gamified evaluation, has shown that such approaches improve both performance and engagement compared to traditional testing (Roig et al., 2026). Furthermore, alternative assessment approaches align closely with contemporary conceptualizations of perceived communicative competence, which emphasize not only linguistic proficiency but also adaptability, contextual awareness, and real-world communication skills (Ertay & Gilanlioglu, 2024). By integrating assessment into authentic classroom activities, these approaches create opportunities for students to practice and refine their communicative abilities in meaningful contexts.

4.2.3 The Association between Student Engagement and Perceived Communicative Competence

The results indicate that student engagement has a strong and significant effect on perceived communicative competence ($\beta = 0.542$), highlighting the importance of active participation and involvement in language learning processes. This finding reinforces the widely accepted view that engagement is a critical determinant of learning outcomes. This result is consistent with prior studies showing that higher levels of engagement are associated with improved academic performance and learning effectiveness. For instance, research on virtual learning environments has demonstrated that limited engagement leads to reduced academic outcomes, while active participation enhances learning success (Rabby et al., 2025). Similarly, studies on AI-supported and interactive learning systems emphasize that engagement—across cognitive, emotional, and behavioral dimensions, plays a key role in shaping learning experiences and outcomes (Mele et al., 2025). In the context of ESL

education, engagement is particularly important because language learning inherently requires interaction, communication, and practice. The present findings suggest that students who are more engaged in classroom activities are better able to develop perceived communicative competence, as they have more opportunities to use language, receive feedback, and refine their skills.

4.3 Implications

4.3.1 Theoretical contributions

This study offers several important theoretical contributions to the literature on ESL education and assessment-driven learning. First, this study advances classroom-based assessment theory by reconceptualizing classroom-based alternative assessment (CLB) as a learning-enabling mechanism rather than a measurement tool. While prior studies have predominantly treated assessment as an evaluative outcome-oriented process, this study demonstrates that assessment practices embedded within classroom activities are associated with learning behaviors and outcomes. By providing evidence of statistically significant relationships on both engagement and perceived communicative competence, the study shifts the theoretical role of assessment from passive evaluation to active learning facilitation. Second, this study contributes to the student engagement literature by extending its role beyond general academic performance into language-specific outcomes, particularly perceived communicative competence. While prior research has acknowledged engagement as a predictor of academic success, limited attention has been given to its role in being associated with perceived communication abilities in ESL contexts. The findings provide empirical evidence that engagement is not only associated with learning participation but also closely linked to the development of communicative capabilities. Third, this study enriches the conceptualization of perceived communicative competence by embedding it within an assessment–learning framework. Existing literature has primarily focused on defining perceived communicative competence in terms of linguistic and intercultural dimensions, with limited exploration of how pedagogical and assessment practices contribute to its development. This study addresses this gap by demonstrating that perceived communicative competence is positively associated with classroom-based assessment practices that emphasize authenticity, interaction, and feedback. Fourth, this study contributes methodologically by applying a prediction-oriented PLS-SEM approach with predictive relevance (Q^2) in ESL research. Compared to traditional explanatory models, this approach highlights the predictive capability of assessment-driven learning models, thereby aligning with the increasing emphasis on predictive analytics in educational research. Finally, by providing empirical evidence from an emerging ESL context, this study extends the generalizability of assessment and engagement theories beyond Western educational settings, offering insights into how assessment practices operate in transitioning educational systems.

4.3.2 Managerial implications

The findings provide several actionable implications for ESL educators, curriculum designers, and university administrators. First, classroom-based alternative assessment should be embedded systematically into course design rather than implemented as occasional supplementary activities. Departments can require each English-related course to include a balanced assessment portfolio comprising diagnostic tasks, formative feedback activities, peer assessment, reflective exercises, and authentic communicative tasks such as presentations, role-plays, simulations, and group discussions. Second, universities can use existing learning management systems, such as Moodle, to operationalize alternative assessment efficiently. For example, the Moodle Quiz function can be used for low-stakes diagnostic assessment, the

assignment function can support audio or video submissions, the workshop module can facilitate structured peer assessment, and discussion forums can capture students' participation in collaborative learning tasks (Appendix A). Digital portfolios can also allow students to document their progress over time by uploading revised written work, presentation recordings, and reflective notes. Third, departments should adopt a standardized communicative assessment rubric to ensure consistency across courses and instructors. The rubric should assess clarity of expression, fluency, linguistic accuracy, interaction, contextual appropriateness, and students' ability to use feedback for improvement (Appendix B). A shared rubric can reduce inconsistency in grading while preserving the authentic and formative nature of alternative assessment. Fourth, institutions should provide targeted professional development rather than general training. Workshops should focus on designing authentic ESL tasks, using Moodle modules for peer review and portfolio assessment, providing rubric-based feedback, and applying AI-supported tools responsibly. Where AI-assisted feedback tools are introduced, institutions should establish clear guidelines regarding transparency, academic integrity, data privacy, and the teacher's role in verifying automated recommendations. Finally, implementation should follow a staged approach. Departments may begin with one or two pilot courses, monitor student participation and teacher workload, refine the rubric and Moodle workflow, and gradually scale the model across the curriculum. This approach can help institutions balance pedagogical innovation with feasibility and quality assurance.

Appendix A and Appendix B are proposed implementation tools derived from the study findings and the broader literature. Their effectiveness should be evaluated through future pilot studies before institution-wide adoption.

5. Conclusion, Limitations and Future research directions

5.1 Conclusion

This study examined the relationships among classroom-based alternative assessment, student engagement, and perceived communicative competence in an English language learning context. The findings indicate that classroom-based alternative assessment is positively associated with student engagement and perceived communicative competence. Student engagement is also positively associated with perceived communicative competence, highlighting the relevance of participation and interaction in language learning environments. These findings suggest that formative, interactive, and authentic assessment practices may support more engaging classroom environments and may be associated with students' confidence in using English for communication. The model demonstrates satisfactory explanatory power and cross-validated predictive relevance. However, because the study relies on cross-sectional survey data and self-reported measures, the findings should not be interpreted as definitive evidence that alternative assessment causes improvements in objectively assessed linguistic performance. The study contributes to the literature by integrating classroom-based alternative assessment, student engagement, and perceived communicative competence within a unified framework. The proposed Moodle-based implementation framework and departmental rubric illustrate how institutions may translate these insights into classroom practice. Future longitudinal, experimental, and mixed-method research should combine students' perceptions with objective performance indicators and explicitly examine technology-mediated assessment practices.

5.2 Limitations and Future research directions

Despite its contributions, this study is subject to several limitations that should be acknowledged. First, the study adopts a cross-sectional research design, which limits the ability

to infer causal relationships among the constructs. Although the findings provide strong evidence of significant associations, future research could employ longitudinal or experimental designs to better capture causal dynamics and changes over time. Second, perceived communicative competence was measured using students' self-reported perceptions. Accordingly, the construct should be interpreted as perceived communicative competence rather than as an objective measure of linguistic performance. Although students' self-evaluations are valuable for understanding confidence, participation, and perceived readiness to communicate, they may not fully reflect demonstrated language proficiency. Future studies should triangulate self-reported perceptions with objective indicators, such as speaking-test scores, presentation grades, teacher evaluations, structured classroom observations, or rubric-based assessments of role-play and discussion performance. Such triangulation would provide a more comprehensive assessment of perceived communicative competence and strengthen construct validity. Third, the study focuses on university students in a specific geographical context (Vietnam), which may limit the generalizability of the findings to other educational settings or cultural contexts. Future research could replicate the model in different countries, educational levels (e.g., secondary education), or institutional environments to validate and extend the findings. Fourth, although the model demonstrates strong explanatory and predictive capabilities, it focuses on a limited set of constructs. Future research could extend the model by incorporating additional variables, such as technological readiness, learning motivation, or instructional design factors, to provide a more comprehensive understanding of ESL learning processes. Finally, given the increasing integration of digital technologies in education, future studies could explore how technology-enhanced assessment practices, such as AI-based feedback systems or online collaborative platforms, interact with classroom-based assessment to influence engagement and perceived communicative competence.

Appendix A: Proposed Moodle integration framework

Assessment stage	Moodle tool	ESL activity	Evidence collected	Feedback method
Diagnostic assessment	Quiz or Questionnaire	Short vocabulary, listening, and self-assessment tasks at the beginning of the course	Baseline scores and learning needs	Automated feedback and teacher guidance
Formative speaking practice	Assignment with audio or video submission	Individual presentation, role-play, or oral reflection	Recorded speaking performance	Teacher rubric and audio feedback
Peer assessment	Workshop module	Students review classmates' presentations or written responses	Peer ratings and qualitative comments	Structured peer-review rubric
Portfolio assessment	Assignment, Database, or e-portfolio link	Students upload revised writing, speaking recordings, and reflective notes	Evidence of progress over time	Periodic teacher feedback
Collaborative discussion	Forum or group activity	Debate, case discussion, or problem-solving task	Participation frequency and response quality	Teacher comments and peer interaction
Reflective learning	Journal, Feedback, or Assignment	Students identify strengths, weaknesses, and improvement goals	Reflective entries	Individualized recommendations
AI-supported extension	Approved AI feedback tool integrated with institutional guidelines	Grammar suggestions, speaking rehearsal, or feedback comparison	Drafts, revisions, and reflection logs	Teacher verification and ethical-use guidance

Appendix B. Proposed Departmental Rubric for Authentic Communicative Assessment

Criterion	Weight	Excellent	Competent	Developing	Needs Improvement
Clarity and organization of ideas	20%	Ideas are consistently clear, logically structured, and easy to follow.	Ideas are generally clear and logically organized, with minor inconsistencies.	Meaning is understandable but organization is uneven or occasionally unclear.	Ideas are difficult to follow because of limited clarity or organization.
Fluency	20%	Communication is sustained smoothly with minimal hesitation.	Communication is generally sustained, with occasional hesitation that does not impede meaning.	Frequent hesitation sometimes interrupts communication.	Hesitation substantially limits communication.
Linguistic accuracy and range	15%	Grammar and vocabulary are accurate, varied,	Minor errors are present but do not	Errors occasionally	Frequent errors substantially

		and appropriate to the task.	interfere with meaning.	affect clarity or limit expression.	impede understanding.
Interaction and responsiveness	20%	Responds appropriately, asks relevant questions, and actively maintains interaction.	Responds appropriately and participates consistently in interaction.	Responds to prompts but shows limited ability to maintain interaction.	Provides minimal responses or struggles to sustain interaction.
Contextual appropriateness	15%	Language is consistently appropriate for the audience, purpose, and context.	Language is generally appropriate, with minor inconsistencies.	Some language choices are unsuitable for the context.	Language choices frequently do not fit the audience or purpose.
Reflection and use of feedback	10%	Demonstrates clear improvement by applying feedback and identifying specific development goals.	Uses feedback appropriately and identifies relevant improvement areas.	Applies feedback inconsistently or identifies only general improvement areas.	Shows limited evidence of using feedback or reflecting on performance.

Recommended application: Departments may embed this rubric into Moodle assignments, workshop peer-review activities, speaking assessments, presentations, simulations, and role-play tasks. Teachers may adapt the descriptors to match course level while retaining the same core assessment dimensions.

References

Ab Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (2017). Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/890/1/012163

Ahmed, T., Kobir, M. H., Yang, Y., Olsen, A. A., Fuad, M., & Deb, S. (2026). Evaluating the efficacy of virtual reality-based training in classroom settings for additive manufacturing education. *Displays*, *91*. https://doi.org/10.1016/j.displa.2025.103257

Badem, F. (2025). Multimodal repair initiations in video-mediated EFL classroom interactions: Focus on screen-based and embodied actions. *Learning, Culture and Social Interaction*, *54*. https://doi.org/10.1016/j.lcsi.2025.100935

Burns, C. S., Digiacomo, D. K., Pusateri, J., Ritchie, K., & Wusylko, C. (2026). Student perspectives on grading in higher education STEM courses. *Information and Learning Sciences*, 1–19. https://doi.org/10.1108/ILS-08-2025-0143

Byers, R. E., Carter, C. R., & Wang, Y. (2024). Student engagement in synchronous online learning: Effectiveness of camera and chat/vote engagement methods. *Decision Sciences Journal of Innovative Education*, *22*(3), 138–157. https://doi.org/10.1111/dsji.12309

Carey, P. (2013). Student engagement: stakeholder perspectives on course representation in university governance. *Studies in Higher Education*, *38*(9), 1290–1304. https://doi.org/10.1080/03075079.2011.621022

Chopra, K. (2025). Beyond the classroom: A retail store based experiential learning approach to business education. *International Journal of Management Education*, *23*(3). https://doi.org/10.1016/j.ijme.2025.101236

Cloudia Ho, Y. Y. (2020). Communicative language teaching and English as a foreign language undergraduates' communicative competence in Tourism English. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *27*. https://doi.org/10.1016/j.jhlste.2020.100271

Ertay, Z., & Gilanlioglu, I. (2024). The Development and Validation of an Intercultural Communicative Competence Scale for Students. *SAGE Open*, *14*(2). https://doi.org/10.1177/21582440241252495

Ferrell, O. C., Harrison, D. E., Ferrell, L., & Hair, J. F. (2019). Business ethics, corporate social responsibility, and brand attitudes: An exploratory study. *Journal of Business Research*, *95*(January), 491–501. https://doi.org/10.1016/j.jbusres.2018.07.039

Fink, M. C., Galler, C., Watzka, B., & Ertl, B. (2026). Effects of annotations and quizzes in a classroom-based immersive virtual field trip. *Computers and Education: X Reality*, *8*. https://doi.org/10.1016/j.cexr.2026.100140

Fonteles, J. H., Cohn, C., Ayalon, E., Zhou, M., T.S., A., Davalos, E., Li, Z., Rayala, S., Mereddy, D., Coursey, A., Jain, S., Zhang, Y., Enyedy, N., Danish, J., & Biswas, G. (2026). Analyzing embodied learning in classroom settings: A human-in-the-loop AI approach for multimodal learning analytics. *Learning and Instruction*, *103*. https://doi.org/10.1016/j.learninstruc.2025.102274

Fuller, M., Kamans, E., van Vuuren, M., Wolfensberger, M., & de Jong, M. D. T. (2021). Conceptualizing Empathy Competence: A Professional Communication Perspective.

Journal of Business and Technical Communication, *35*(3), 333–368. https://doi.org/10.1177/10506519211001125

Guilloteaux, M. J. (2013). Motivational strategies for the language classroom: Perceptions of Korean secondary school English teachers. *System*. https://doi.org/10.1016/j.system.2012.12.002

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. In *European Business Review* (Vol. 31, Number 1, pp. 2–24). Emerald Group Publishing Ltd. https://doi.org/10.1108/EBR-11-2018-0203

Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). University of Chicago Press.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Koponen, J., Turunen, M., Laajalahti, A., Julkunen, S., & Spitzberg, B. H. (2025). Managers' Technology-Mediated Communication Competence: A Theoretical Framework. *Journal of Business and Technical Communication*, *39*(3), 338–397. https://doi.org/10.1177/10506519251326558

Kozaklı Ülger, T., Kıryak, Z., & Ülger, B. B. (2026). Teachers' noticing and social interactions in the classroom: Insights from a skill-based science curriculum. *Thinking Skills and Creativity*, *61*. https://doi.org/10.1016/j.tsc.2026.102225

Lee, S., & Jeon, J. (2024). Teacher agency and ICT affordances in classroom-based language assessment: The return to face-to-face classes after online teaching. *System*, *121*. https://doi.org/10.1016/j.system.2023.103218

Mele, C., Russo-Spena, T., Ranieri, A., Di Bernardo, I., & Hollebeek, L. D. (2025). Talking vs typing: how voice- vs textbased educational robots shape student engagement. *Journal of Service Management*. https://doi.org/10.1108/JOSM-02-2024-0082

Muhonen, H., Pakarinen, E., & Lerkkanen, M. K. (2025). Is teachers' professional vision visible in the classroom? Exploring the links between teachers' knowledge-based reasoning and observed classroom interactions. *Learning and Instruction*, *100*. https://doi.org/10.1016/j.learninstruc.2025.102220

O'Brien, A., Clavin, A., Horkan, S., Cassidy, A., Duane, S., Cronin, N., Mitchell, S., Clancy, J., Horan, L. A., Sammon, J., & Kindermann, G. (2026). Co-creating sustainable learning communities in a transdisciplinary course. *Information and Learning Sciences*, 1–20. https://doi.org/10.1108/ILS-09-2025-0172

Priante, A., & Tsekouras, D. (2025). Integrating technology in physical classrooms: The impact of game-based response systems on student learning experience. *Information and Management*, *62*(3). https://doi.org/10.1016/j.im.2025.104105

Rabby, S. M., Akhi, A. A., Talukder, M. S., & Das, S. (2025). How does limited student engagement in virtual classrooms affect students' academic performance? A stressor–strain–outcome perspective. *On the Horizon*, *33*(3), 357–377. https://doi.org/10.1108/OTH-02-2025-0017

Rahmanu, I. W. E. D., & Molnár, G. (2026). EFL classroom settings: Behavioral intention of using screen-based 360-degree video in learning English language. *Social Sciences and Humanities Open*, *13*. https://doi.org/10.1016/j.ssaho.2026.102576

Roig, P. J., Alcaraz, S., Gilly, K., Bernad, C., & Juiz, C. (2026). A multistage linear path design for educational escape rooms. *International Journal of Comparative Education and Development*, 1–21. https://doi.org/10.1108/IJCED-09-2025-0121

Sahlender, M., Helbert, S., ten Hagen, I., Knaus, A., & Weiss, Z. (2024). Investigating spoken classroom interactions in linguistically heterogeneous learning groups – An interdisciplinary approach to process video-based data in second language acquisition classrooms. *Applied Corpus Linguistics*, *4*(3). https://doi.org/10.1016/j.acorp.2024.100104

Sanford, D., & Snow, A. (2026). Beyond the lecture: Using formative assignments to overcome classroom challenges. *Journal of Professional Nursing*, *62*, 92–97. https://doi.org/10.1016/j.profnurs.2025.11.001

Sarwari, A. Q., Adnan, H. M., Rahamad, M. S., & Abdul Wahab, M. N. (2024). The Requirements and Importance of Intercultural Communication Competence in the 21st Century. In *SAGE Open* (Vol. 14, Number 2). SAGE Publications Inc. https://doi.org/10.1177/21582440241243119

Sieveke, P., Graham, S., Souvignier, E., & Busse, V. (2026). From research to classroom: A mixed-methods investigation into implementing evidence-based writing practices. *Teaching and Teacher Education*, *177*. https://doi.org/10.1016/j.tate.2026.105519

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, *36*(2), 111–147. http://www.jstor.org/stable/2984809

Thi Viet, D. D., & Nguyen, L.-T. (2025). Unveiling the Impact of Big Data and Predictive Analytics Adoption on Sustainable Supply Chain Management: An Employee-Centric Perspective. *SAGE Open*, *15*(3). https://doi.org/10.1177/21582440251363128

Unesco. (2025). *Record number of higher education students highlights global need for recognition of qualifications*. Unesco. https://www.unesco.org/en/articles/record-number-higher-education-students-highlights-global-need-recognition-qualifications

Xu, J., Tong, M., Gao, Q., Xu, Y., Chen, J., & Long, T. (2026). The Impact of project-based learning (PjBL) on creative thinking in fourth graders: A quasi-experimental study in blended science synchronous classroom. *Thinking Skills and Creativity*, *61*. https://doi.org/10.1016/j.tsc.2026.102165

Yu, J., Li, S., Liu, H., Liu, Q., Tan, C., Cheng, Z., & Wu, J. (2026). EduYOLO: A classroom behavior recognition framework based on high-resolution feature attention fusion. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2026.131370

#197 HUFLIT ENGLISH MAJORS' PERCEPTIONS OF AI-SUPPORTED TRANSLATION LEARNING IN A VIETNAMESE EFL CONTEXT

Nam Anh Tran[1], Thanh Tu Tran[2]

Abstract

The rapid integration of generative AI tools such as ChatGPT and Grammarly is reshaping EFL translation instruction in Vietnamese higher education, yet empirical evidence from specific institutional contexts remains limited. This study investigates HUFLIT English majors' perceptions of AI-supported translation learning through an integrated Technology Acceptance Model (TAM), Technological Pedagogical Content Knowledge (TPACK), and Self-Regulated Learning (SRL) framework. Using a structured questionnaire administered to 219 second- to fourth-year English majors, the study examined perceived benefits and risks of AI, teacher AI integration, and metacognitive engagement during translation tasks. Descriptive statistics revealed high AI adoption rates (92.7%), strong perceived usefulness and ease of use, and widespread self-reported cognitive overreliance. Exploratory factor analysis (KMO = 0.899; four factors explaining 73.53% of variance) identified a four-factor model: AI Learning Engagement, Teacher AI Integration, Cognitive Overreliance, and Perceived Ease of Use. Notably, perceived usefulness and self-regulated learning converged into a single construct, suggesting that HUFLIT students perceive AI as valuable precisely because it scaffolds their metacognitive processes. These findings underscore the critical role of teacher guidance and institutional support in supporting responsible, critically engaged AI use in Vietnamese EFL translation programs.

Keywords: Artificial Intelligence in translation learning, English majors' perceptions, translation pedagogy, self-regulated learning, Vietnamese higher education

1. Introduction

1.1. Background of the Study

The rapid development of artificial intelligence (AI), particularly generative AI tools such as ChatGPT and Grammarly, is reshaping both language education and the translation profession. Professional translators today increasingly function as post-editors who evaluate and refine machine-generated output rather than producing translations entirely from scratch (Ren & Wang, 2025; Briggs, 2018). As a result, translation competence now requires not only linguistic proficiency but also the ability to critically evaluate AI-generated text and use automated support strategically.

In higher education, however, translation teaching does not always keep pace with these professional changes. Many translation programmes still rely on teacher-centred instruction and conventional task designs, with limited systematic integration of AI-supported workflows (Yan et al., 2025; Sutrisno et al., 2024). In Vietnamese higher education, this gap is particularly significant: English learners often depend heavily on formal classroom instruction and may

[1] *Corresponding author, Lecturer, Faculty of Foreign Languages, HCMC University of Foreign Languages – Information Technology, Ho Chi Minh City, Vietnam; Email: anhtn@huflit.edu.vn*
[2] *Lecturer, Faculty of Foreign Languages, HCMC University of Foreign Languages – Information Technology, Ho Chi Minh City, Vietnam*

have fewer opportunities for authentic English exposure outside class. At HUFLIT, English majors in EFL translation courses are increasingly exposed to AI tools such as ChatGPT and Grammarly, yet how they perceive the benefits, risks, and pedagogical value of such tools has not been sufficiently examined.

To address this, the present study draws on three complementary theoretical frameworks. The Technology Acceptance Model (TAM; Davis, 1989) explains how learners' perceptions of usefulness and ease of use shape their adoption of educational technology. Technological Pedagogical Content Knowledge (TPACK; Mishra & Koehler, 2006) examines how teachers integrate technology with pedagogical and content knowledge in instructional practice. Self-Regulated Learning (SRL; Zimmerman, 2000) highlights learners' metacognitive processes – including planning, monitoring, and reflection, when engaging with learning tools. Together, these frameworks provide a multidimensional lens for understanding AI-supported translation learning.

1.2. Research Problem and Gap

Although AI has become an increasingly prominent topic in English language education, much of the existing research focuses on general ESL or EFL learning rather than the specific demands of translation instruction. Translation learning requires students not only to understand language form but also to compare source and target texts, evaluate semantic accuracy, consider stylistic appropriateness, and justify their translation choices. Findings from broader language-learning studies may therefore not fully capture how students experience AI in translation classrooms.

In addition, empirical evidence from Vietnamese higher education remains limited, particularly regarding how English majors perceive AI-supported translation learning in specific institutional settings. Existing studies also tend to rely on single theoretical lenses – such as technology acceptance or pedagogical integration alone, which are insufficient for capturing the interaction among students' perceptions of AI, teachers' instructional practices, and learners' self-regulated engagement (Manzoor et al., 2025; Saud, 2025). This study addresses these gaps by investigating HUFLIT English majors' perceptions through an integrated TAM–TPACK–SRL framework.

1.3. Research Objectives

This study aims to investigate HUFLIT English majors' perceptions of AI-supported EFL translation learning through an integrated TAM–TPACK–SRL framework. Specifically, the study examines students' perceived benefits and risks of AI use, their perceptions of teacher AI integration in translation classes, and the empirical factor structure underlying these perceptions.

1.4. Research Questions

The study is guided by the following research questions:

1. How do HUFLIT students perceive the benefits and risks of AI-supported translation learning?

2. What empirical factor structure emerges from students' perceptions of AI-supported translation learning?

3. How do students perceive the roles of teacher AI integration and self-regulated engagement when using AI in translation classes?

1.5. Significance of the Study

This study contributes to the literature on AI-supported language learning by providing empirical evidence from a Vietnamese higher education context that has been underexplored in existing research. Theoretically, it examines how an integrated TAM–TPACK–SRL framework applies to translation-specific instruction, and empirically identifies a four-factor perceptual model of AI-supported translation learning. Practically, the findings offer actionable insights for teachers and institutions seeking to design more effective AI-supported translation curricula, promote responsible AI use, and develop context-sensitive policies for English majors in Vietnam.

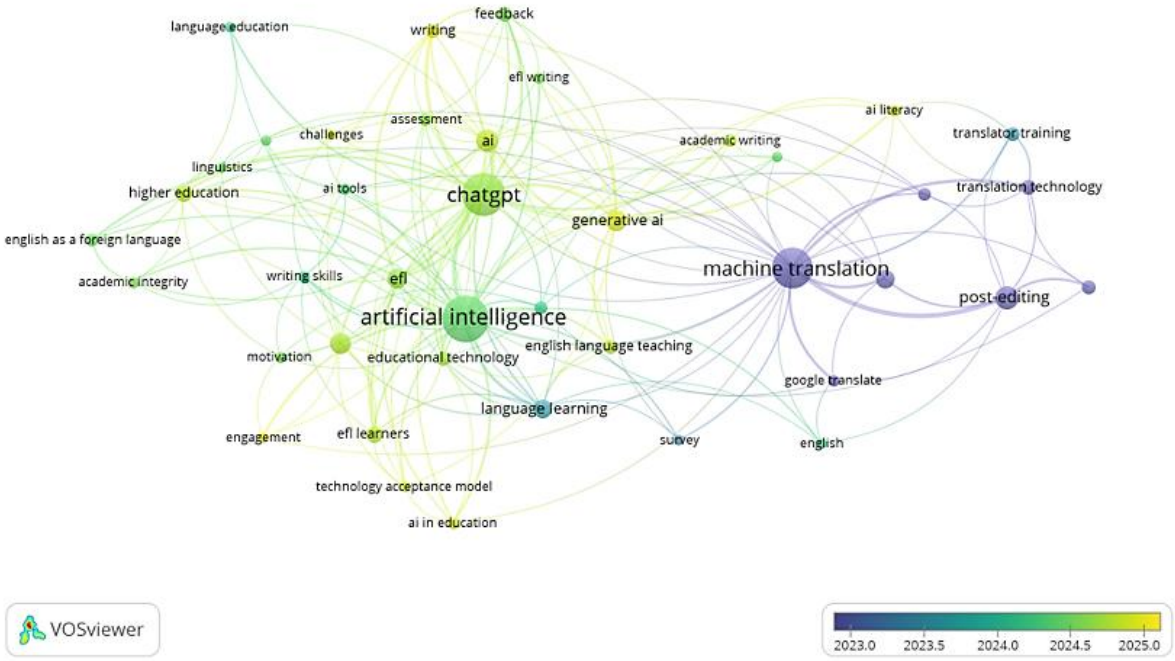
2. Literature Review

This chapter reviews prior research on the integration of artificial intelligence (AI) into English language education and translation learning, with particular attention to two dimensions central to the present study: pedagogical effectiveness and ethical considerations. In recent years, generative AI tools such as ChatGPT, grammar checkers, and machine translation systems have become increasingly integrated into language learning environments. Understanding how these tools influence students' learning experiences and perceptions is therefore essential for evaluating their pedagogical potential.

2.1. AI in English Language Education

AI has rapidly become a salient feature of contemporary English language education, shifting traditional teacher-fronted classrooms toward more learner-centered, technology-rich environments. AI tools not only provide on-demand language feedback but also facilitate crucial out-of-class practice and distributed learning (Saud, 2025; Manzoor et al., 2025; Rahman, 2026). While general EFL learning often uses AI for idea generation or basic grammar correction, translation classrooms involve different cognitive demands.

Figure 8
Analysis of recent literature



To map the current research landscape, a bibliometric analysis was conducted using VOSviewer. Publication data were retrieved from the Scopus database on [30 June 2026], using the following search query: TITLE-ABS-KEY ((“artificial intelligence” OR “ChatGPT” OR “generative AI” OR “machine translation”) AND (“language learning” OR “EFL” OR “ESL” OR “English language teaching” OR “foreign language” OR “translation”) AND (“perception” OR “attitude” OR “acceptance” OR “student” OR “learner”)). The search was restricted to journal articles (DOCTYPE, “ar”) published between 2018 and 2025 (PUBYEAR > 2017 AND PUBYEAR < 2026), written in English, classified under the Social Sciences (SOCI) or Arts and Humanities (ARTS) subject areas, and available in open access. The search returned 426 articles, which were exported and analysed in VOSviewer. A co-occurrence analysis of author keywords was performed using full counting, with a minimum threshold of five occurrences per keyword, resulting in 44 keywords that met this criterion. The resulting network visualisation (Figure 1) reveals the following patterns:

- The network was organised into three major thematic clusters. The first cluster centred on “machine translation”, “post-editing”, “neural machine translation”, and “translator training”, representing the traditional translation-technology strand. The second cluster, organised around “artificial intelligence”, “ChatGPT”, and “generative AI”, reflected the more recent shift toward generative AI tools. The third cluster grouped pedagogical and learner-oriented terms such as “EFL”, “language learning”, “writing”, “feedback”, and “technology acceptance model”.

- The overlay visualisation, which maps keywords by their average publication year, indicated a clear temporal shift. Specifically, earlier research clustered around “machine

translation" and "Google Translate", whereas more recent work has moved toward "generative AI", "ChatGPT", and "technology acceptance".

- Notably, the links between the translation-instruction cluster and learner-perception terms remained relatively weak, suggesting that the application of AI specifically within translation pedagogy, as distinct from general language learning, has not yet been fully explored.

Taken together, this bibliometric evidence supports the integrated TAM–TPACK–SRL framework adopted in the present study and highlights the underexplored intersection of AI and EFL translation instruction, which this study seeks to address.

Students must compare source and target texts, evaluate semantic accuracy, consider stylistic appropriateness, and justify their renderings. Thus, AI in translation learning serves not only as a language improver but also as a cognitive scaffold for text analysis and post-editing, requiring a distinct set of pedagogical and self-regulatory skills.

2.2. Pedagogical Effectiveness of AI-Supported Learning

2.2.1. Language support: grammar, vocabulary, and writing

A substantial body of work has explored AI's potential to enhance core language skills, particularly in grammar, vocabulary, and writing. Rahman (2026) review in ETDC highlights that tools such as Grammarly, Quillbot, and ChatGPT can help learners generate ideas, improve lexical choice, and reduce grammatical errors in drafts. Learners tend to perceive AI-based grammar correction tools as helpful for identifying mistakes they might otherwise overlook and for increasing the accuracy and fluency of their written output. In their systematic review, Manzoor et al. (2025) report that repeated interaction with AI feedback can be associated with reductions in certain error types, provided that students engage reflectively with the suggestions rather than accepting them mechanically. Similar attitudinal patterns have been reported in translation-classroom contexts more broadly, where students value machine translation as a support tool while remaining cautious about its accuracy (Pastor, 2021).

In writing, AI-supported tools can serve as both drafting and revision aids. A quasi-experimental study on AI-assisted learning and EFL writing proficiency shows that students who received AI-mediated feedback improved more in overall writing scores than those who relied solely on traditional feedback (Alangari, 2025). AI feedback is often perceived as faster and more immediate than traditional teacher feedback, although teachers may provide more nuanced explanations and contextual guidance (Alangari, 2025). However, both Manzoor et al. (2025) and Saud (2025) caution that the pedagogical value of AI feedback depends heavily on how learners interact with it. When students use AI suggestions as prompts for self-explanation, comparison, and revision, the tools can support deeper language awareness; when they accept suggestions passively, AI risks becoming a shortcut rather than a learning opportunity. For translation learners, such feedback may support the development of error awareness and lexical sensitivity when students compare AI-generated suggestions with their own translations.

2.2.2. Self-paced learning and learner autonomy

Beyond discrete language skills, AI has been promoted as a means to foster self-paced learning and learner autonomy. Manzoor et al. (2025) describe how AI-integrated systems can adapt task difficulty to learners' performance, provide tailored recommendations, and support flexible practice schedules. Saud (2025) add that conversational agents and chatbots can offer extended opportunities for language use outside the classroom, which is particularly relevant in EFL contexts where authentic interaction is limited. When learners perceive AI tools as

controllable, useful, and aligned with their goals, they are more likely to integrate them into their personal learning routines as supplemental tutoring.

At the same time, promoting autonomy through AI is not automatic. The ERIC paper on ethical issues notes that some students become overly dependent on AI for answers and corrections, especially when they are under time pressure or lack confidence in their language ability (Nhan et al., 2025). In such cases, AI can reduce opportunities for productive struggle and experimentation, which are important for long-term language development. These findings suggest that teachers need to frame AI use explicitly as a support for learning processes rather than as a shortcut to finished products, for example by asking students to justify why they accept or reject particular AI suggestions. These characteristics align with the concept of self-regulated learning, where learners actively monitor their progress, evaluate feedback, and adjust their learning strategies during task completion (Zimmerman, 2000).

2.2.3. Risks of overreliance and passive learning

An emerging strand of research focuses on potential negative effects of AI-supported learning on learners' cognitive and metacognitive development. The ERIC article emphasizes that overreliance on AI can lead to offloading not only mechanical aspects of writing but also higher-order thinking, such as organizing ideas and evaluating information (Nhan et al., 2025). When learners rely on AI-generated text or suggestions without critical evaluation, they may gradually lose confidence in their own linguistic and rhetorical abilities. Manzoor et al. (2025) similarly argue that passive acceptance of AI feedback can interfere with the development of metalinguistic awareness, as students focus on surface correctness rather than underlying grammatical or discourse patterns.

These concerns do not negate AI's pedagogical value, but they underscore the need for more nuanced integration. Manzoor et al. (2025) recommend that AI tools be embedded in instructional designs that promote reflection, questioning, and comparison between AI output and students' original work. In a similar vein, Compagnoni et al.'s (2025) article calls for cultivating critical AI literacy so that learners evaluate AI output rather than accepting it uncritically. For English majors who are likely to become language professionals or educators, developing such critical literacy is especially important.

2.3. Ethical Considerations in AI-Supported Learning

2.3.1. Academic integrity, authorship, and authenticity

Ethical debates around AI in education have largely centered on academic integrity and the boundaries between legitimate support and academic misconduct. In language education, these concerns are particularly salient because AI systems can generate fluent, human-like text on demand. The ERIC paper documents how students and teachers struggle to define when AI-assisted writing becomes plagiarism and how authorship should be understood when AI contributes to idea generation, organization, or phrasing (Nhan et al., 2025). Some participants viewed AI use for surface-level editing as acceptable, while considering extensive reliance on AI-generated content without acknowledgment to be unethical. These ethical issues are particularly relevant in translation learning, where students may rely on AI-generated translations or paraphrases during coursework and assessments.

Compagnoni et al. (2025) also points out that institutional policies often lag behind technological developments, resulting in inconsistent expectations about AI use across courses and institutions. Policy ambiguity can create anxiety for learners who wish to use AI responsibly but are unsure where the line is drawn, especially in high-stakes assignments. For

English majors, whose coursework typically involves substantial written production and critical engagement with texts, clarity about what constitutes ethical AI use is crucial.

2.3.2. Data privacy, fairness, and transparency

Another cluster of ethical concerns relates to data privacy, fairness, and transparency in AI systems. Many AI tools collect and process user data, including text input and usage patterns, yet learners may not fully understand how their data are stored, who can access them, or how they may be used to further train commercial models. Compagnoni et al. (2025) emphasizes that transparency and accountability are key principles for ethical AI in language pedagogy and argues that teachers and institutions should help students understand both the capabilities and limitations of AI tools they are expected to use.

Fairness and bias also present significant challenges. Because AI systems draw on large training datasets, they can reproduce linguistic and cultural biases, for instance by privileging certain varieties of English or by providing feedback that implicitly enforces narrow norms of correctness. Manzoor et al. (2025) note that such biases may undermine learners' sense of legitimacy, particularly in contexts where local varieties of English and multilingual practices are part of everyday communication. In response, both the systematic review and the ethical frontiers article call for more context-sensitive evaluation of AI tools and for pedagogical approaches that encourage learners to question unexpected or problematic AI suggestions rather than accepting them uncritically.

2.3.3. Policy responses and contextualized guidelines

In response to these ethical challenges, institutions and professional organizations have begun to articulate principles for responsible AI use in education. Compagnoni et al. (2025) proposes a framework centered on transparency, accountability, and equity, arguing that AI should remain under human control and that learners should be explicitly taught how to use AI ethically, including when and how to disclose AI assistance in academic work. The ERIC article likewise recommends that universities provide clear, discipline-specific guidelines on AI use in coursework, rather than blanket bans or vague statements (Nhan et al., 2025).

Saud (2025) observe that policy responses vary widely across ESL/EFL contexts, with some institutions restricting AI use in high-stakes assessments, while others encourage experimentation combined with explicit training and reflection. Both Saud (2025) and Manzoor et al. (2025) highlight the importance of grounding such policies in empirical evidence about how students actually use AI and what they perceive as benefits and risks. This suggests a need for studies that capture local perceptions and that can inform contextually appropriate guidelines, particularly in rapidly changing contexts such as Vietnam.

2.4. Summary and Research Gap

Existing literature demonstrates that AI-supported learning offers meaningful pedagogical benefits, particularly in vocabulary, writing support, and self-paced learning (Manzoor et al., 2025), while simultaneously presenting risks related to cognitive overreliance, passive learning, and academic integrity (Nhan et al., 2025; Compagnoni et al., 2025). Despite these insights, critical gaps remain. First, most studies focus on general ESL/EFL contexts rather than the specific cognitive demands of translation instruction. Second, empirical research in Vietnamese higher education remains scarce, leaving institutional factors underexplored. Finally, previous studies tend to adopt single frameworks, failing to capture the multidimensional dynamics of EFL translation. Addressing these gaps requires a more

comprehensive approach that examines how students' technology acceptance, teachers' pedagogical practices, and learners' self-regulation intersect in AI-supported translation tasks.

2.5. Theoretical Framework

This study integrates Technology Acceptance Model (TAM) (Davis, 1989), Technological Pedagogical Content Knowledge (TPACK) (Mishra & Koehler, 2006), and Self-Regulated Learning (SRL) (Zimmerman, 2000) to examine English majors' perceptions of AI-supported EFL translation classes at HUFLIT.

TAM provides the core mechanism: perceived usefulness (learning benefits: error awareness, vocabulary expansion from Sections 2.2.1–2.2.2) and perceived ease of use (faster AI vs. detailed teacher feedback) predict intentions to use AI tools. In EFL contexts, TAM explains technology adoption patterns amid limited exposure, but overlooks pedagogical integration and self-regulation processes. In the context of AI-supported translation learning, perceived usefulness may relate to students' beliefs that AI tools improve translation accuracy, vocabulary selection, or error awareness.

TPACK addresses this by framing AI-supported teaching practices as intersections of:

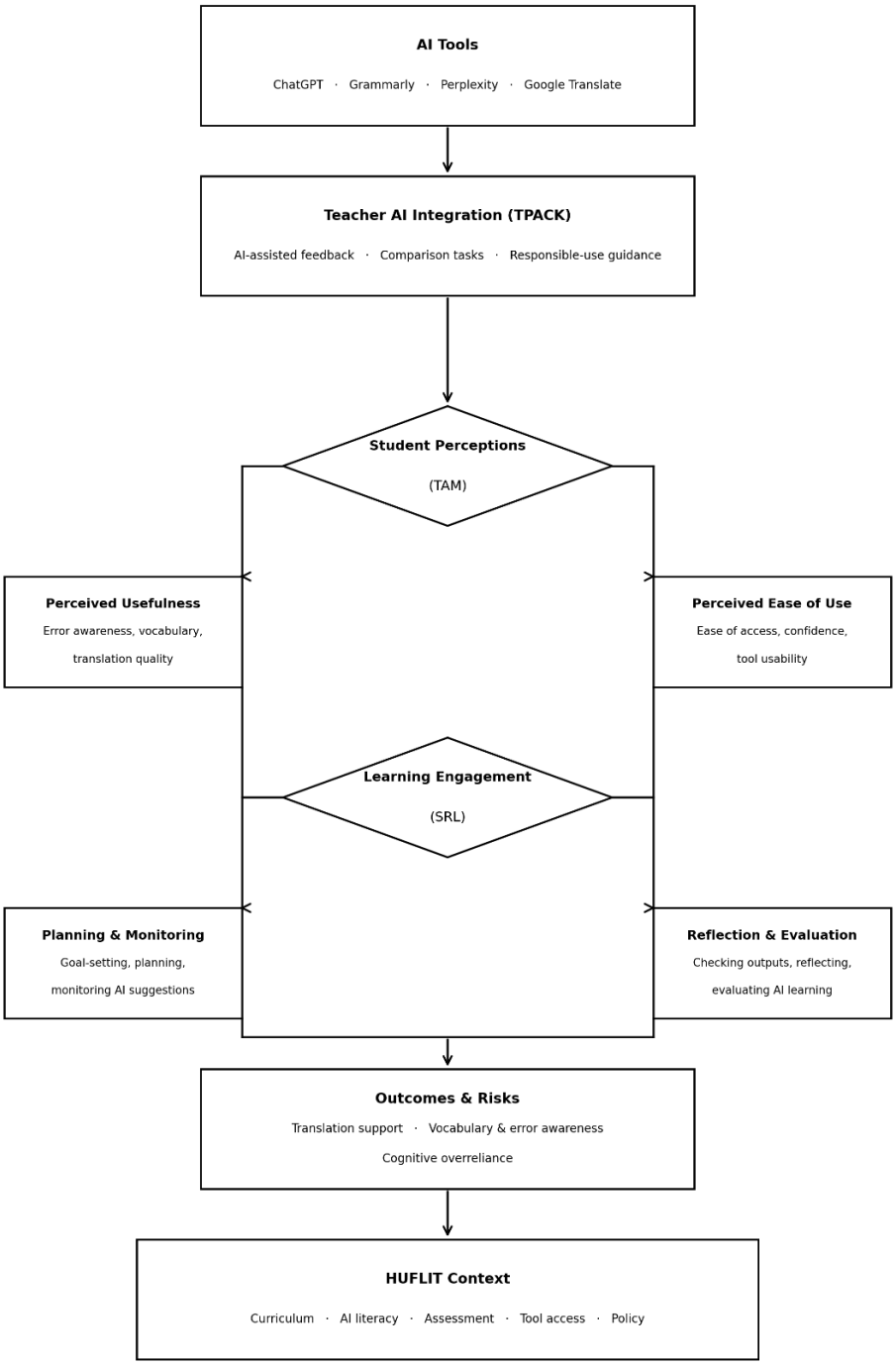
- Technological Knowledge (TK): ChatGPT/Grammarly for adaptive tasks, grammar analysis
- Pedagogical Knowledge (PK): Reduced teacher workload via automated feedback
- Content Knowledge (CK): EFL translation/grammar competencies

Within translation classrooms, TPACK highlights how teachers design tasks that integrate AI tools into pedagogically meaningful activities such as translation comparison, error analysis, and post-editing exercises.

SRL explains metacognitive outcomes, where AI scaffolds forethought (goal-setting), performance (feedback monitoring), and self-reflection phases. Overreliance however risks "cognitive offloading" without guidance. Through these processes, learners may develop metacognitive awareness as they evaluate AI suggestions, revise their translations, and reflect on alternative linguistic choices.

Integration: TAM captures perceptions → TPACK structures teaching → SRL drives learning outcomes. HUFLIT moderators (curriculum/assessments, AI policies/literacy, tool access) contextualize this in Vietnam EFL.

Figure 9
Integrated TAM–TPACK–SRL Framework for AI-Supported Translation Learning



Note. Figure 2 illustrates this integrated model, positioning student perceptions as the core mechanism linking AI teaching practices to learning outcomes, moderated by the HUFLIT context.

3. Methodology

3.1. Research Design

This study adopted a quantitative survey design to investigate HUFLIT English majors' perceptions of AI-supported EFL translation learning. A quantitative approach was appropriate because the study aimed to measure students' perceived benefits and risks of AI, their views of teacher AI integration, and their self-regulated learning behaviours in a systematic and scalable way. The study was guided by an integrated theoretical framework combining the Technology Acceptance Model (TAM), Technological Pedagogical Content Knowledge (TPACK), and Self-Regulated Learning (SRL), which informed both questionnaire construction and the interpretation of results. Within this framework, TAM was used to examine perceived usefulness and perceived ease of use of AI tools, TPACK was used to capture students' perceptions of teachers' AI-supported instructional practices, and SRL was used to explore students' metacognitive engagement when working with AI during translation tasks.

3.2. Participants

The participants were English majors at HUFLIT who had experience with EFL translation courses. Students from the second, third, and fourth years of the undergraduate programme were targeted because they were more likely to have engaged in translation-related coursework and had prior exposure to AI tools in their learning. This participant group was considered appropriate because they could provide informed and experience-based perceptions of AI-supported translation instruction.

3.3. Research Instrument

Data were collected using a structured self-report questionnaire developed on the basis of the TAM–TPACK–SRL framework and prior AI-in-education studies. The instrument consisted of two main parts.

The first part gathered background information, including gender, year of study, frequency of AI tool use in translation learning, prior use of AI in translation courses, main purposes of AI use (translation, grammar checking, vocabulary, idea generation), and the AI tool used most frequently.

The second part measured students' perceptions using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Items were designed to represent five theoretical constructs: Perceived Usefulness (PU, 4 items), Perceived Ease of Use (PEOU, 4 items), Teacher AI Integration (TAI, 5 items), Self-Regulated Learning (SRL, 6 items), and Concerns and Overdependence (CO, 4 items), yielding a total of 23 Likert-scale items. These items were subsequently examined through reliability analysis and exploratory factor analysis to determine the underlying empirical factor structure of the questionnaire.

3.4. Data Collection Procedure

The questionnaire was administered to English majors who were currently taking or had recently completed translation courses at HUFLIT. Participation was voluntary, and students were informed about the purpose of the study and assured that their responses would be used solely for academic research. After collection, the responses were screened and prepared for statistical analysis.

3.5. Data Analysis

The data were analysed using SPSS in four stages. First, descriptive statistics (frequencies, percentages, means, and standard deviations) were used to summarise participants' demographic characteristics and AI usage patterns. Second, descriptive statistics were calculated for all Likert-scale items to describe students' perceptions across the five

theoretical constructs. Third, reliability analysis using Cronbach's alpha was conducted to assess the internal consistency of each scale. Finally, exploratory factor analysis (EFA) using Principal Component Analysis (PCA) with Varimax rotation was performed to examine the underlying empirical structure of the questionnaire and determine whether the theoretical constructs held empirically in this context.

4. Results

This chapter presents the findings of the study in four parts: respondents' demographic profile and AI usage patterns, descriptive statistics of students' perceptions, reliability analysis, and exploratory factor analysis of the measurement instrument.

4.1. Reliability Analysis

The reliability analysis pointed to strong internal consistency across all five initial scales. As shown in Table 1, the Cronbach's alpha values ranged from 0.884 to 0.923, all of which lie well above the commonly accepted threshold of 0.70. In addition, the corrected item–total correlations ranged from 0.658 to 0.875, and, on this basis, no items had to be removed prior to the subsequent analyses.

Table 48
Reliability Analysis Results

Construct	Items	Cronbach's Alpha	Interpretation
Perceived Usefulness (PU)	4	0.884	Good
Perceived Ease of Use (PEOU)	4	0.895	Good
Teacher AI Integration (TAI)	5	0.917	Excellent
Self-Regulated Learning (SRL)	6	0.923	Excellent
Concerns and Overdependence (CO)	4	0.885	Good

4.2. Demographic Profile and AI Usage Patterns

A total of 219 English majors at HUFLIT completed the questionnaire with no missing responses. The sample was predominantly female (78.5%, n = 172), while male students accounted for 21.5% (n = 47). In terms of academic year, most respondents were second-year students (84.0%, n = 184), followed by third-year students (15.1%, n = 33) and fourth-year students (0.9%, n = 2).

Table 49
Demographic Profile of Respondents

Variable	Category	Frequency	Percentage
Gender	Male	47	21.5%

Variable	Category	Frequency	Percentage
	Female	172	78.5%
Year of study	Second year	184	84.0%
	Third year	33	15.1%
	Fourth year	2	0.9%
Total		219	100.0%

AI use in translation learning appeared to be both widespread and frequent. On a five-point frequency scale, the mean score for AI use reached 3.84, with 42.5% of students reporting that they used AI “often” and a further 26.0% “very often”. Only a small minority reported using AI “rarely” or “never”. In line with this pattern, 92.7% of respondents indicated that they had already used AI tools in their translation courses, which suggests that AI has become an integral part of students’ translation learning practices.

As for the purposes of use, grammar checking was the most common (32.9%), followed by vocabulary search (24.2%), idea generation (23.7%), and translation tasks themselves (19.2%). This distribution implies that students tended to rely on AI primarily as a support tool for language improvement, rather than as a full substitute for their own translation work. With regard to platforms, ChatGPT was by far the most frequently used tool (72.6%), whereas Grammarly, Google Translate, and Perplexity were each selected by considerably smaller proportions of respondents.

Table 50
AI Usage Patterns

Variable	Category	Frequency	Percentage
AI use frequency	Never	3	1.4%
	Rarely	16	7.3%
	Sometimes	50	22.8%
	Often	93	42.5%
	Very often	57	26.0%
Main purpose	Grammar checking	72	32.9%

Variable	Category	Frequency	Percentage
	Vocabulary search	53	24.2%
	Idea generation	52	23.7%
	Translation	42	19.2%
Tool used most	ChatGPT	159	72.6%
	Grammarly	34	15.5%
	Google Translate	19	8.7%
	Perplexity	7	3.2%

Note. The "Main purpose" variable was a single-choice item, meaning each respondent selected only one primary purpose. Therefore, the total frequency equals 219 and the total percentage equals 100%. All frequency and percentage values in this table are based on the total sample of N = 219.

4.3. Students' Perceptions of AI-Supported Translation Learning

The descriptive statistics for the Likert-scale items indicate that, overall, students held positive perceptions across all of the main dimensions of AI-supported translation learning, although the strength of these perceptions varied from one construct to another.

Perceived Usefulness recorded the highest ratings of all constructs, with mean scores ranging narrowly from 4.15 to 4.16. Students broadly agreed that AI helped them improve translation quality, acquire vocabulary and appropriate expressions, identify grammatical and lexical errors, and complete translation tasks more effectively. Taken together, these consistently high means suggest that students perceived AI as a genuinely beneficial resource in their translation learning.

Turning to Perceived Ease of Use, the means were slightly lower but still high, ranging from 3.93 to 4.15. In general, students agreed that AI tools were easy to operate and that they could learn to use them for translation tasks without much difficulty. Notably, the item "I feel confident when using AI in translation learning" had the lowest mean in this group (M = 3.93), suggesting that while AI is broadly perceived as accessible, a proportion of students may still benefit from additional guidance or training to fully build their confidence.

Ratings for Teacher AI Integration spanned a somewhat wider range, from 3.96 to 4.33. Students reported that their teachers explained AI use clearly, incorporated AI appropriately into lessons, encouraged translation-comparison activities, and advised them on when to rely on AI and when to avoid it. The highest-rated item in the whole questionnaire, "My teacher helps students understand when they should and should not rely on AI" (M = 4.33), underscores the central role of teacher mediation in shaping responsible AI use.

Self-Regulated Learning also attracted high ratings, with means ranging from 3.90 to 4.17. Across the items, students reported setting goals before using AI, planning their use of AI for particular tasks, monitoring whether AI suggestions were appropriate, checking AI-

generated outputs before accepting them, reflecting after use, and evaluating what AI had contributed to their learning. These responses indicate that many students engaged with AI in ways that support metacognitive awareness, rather than relying on it for purely passive copying.

By contrast, Concerns and Overdependence produced more moderate and more varied ratings, with means ranging from 3.59 to 4.03. The item "I sometimes depend too much on AI when doing translation tasks" received the highest mean in this group (M = 4.03), which points to a fairly widespread awareness of over-reliance. Items concerning reduced independent thinking and the uncritical acceptance of AI suggestions also recorded moderately high means; however, their larger standard deviations indicate more divided responses among students. Taken together, the descriptive results reveal a consistent dual pattern: students clearly value AI's pedagogical benefits and ease of access, yet at the same time they recognise tangible risks associated with dependence and reduced cognitive effort.

4.4. Reliability and Factor Structure

The exploratory factor analysis extracted four factors, which together accounted for 73.526% of the total variance, a figure that comfortably exceeds the recommended 50% threshold. In the rotated component matrix, each item loaded clearly onto a single factor, with no substantial cross-loadings. As summarised in Table 4, however, the resulting structure departed from the initial five-construct model in one important respect: the items originally designed to measure Perceived Usefulness and Self-Regulated Learning converged onto a single factor, which was labelled AI Learning Engagement (ALE). The remaining three factors – Teacher AI Integration (TAI), Cognitive Overreliance (CO), and Perceived Ease of Use (PEOU) – retained the individual structures originally hypothesised.

Table 51
Empirical Factor Structure After EFA

Factor	Label	Items	Variance Explained (rotated)	Original Constructs
F1	AI Learning Engagement (ALE)	10	26.373%	PU + SRL merged
F2	Teacher AI Integration (TAI)	5	17.021%	TAI retained
F3	Cognitive Overreliance (CO)	4	15.407%	CO retained
F4	Perceived Ease of Use (PEOU)	4	14.725%	PEOU retained
Total	**4 factors**	**23**	**73.526%**	5 → 4 constructs

The emergence of AI Learning Engagement (ALE) as a composite factor is the most theoretically significant result of this study. More specifically, ALE refers to a combined pattern in which students perceive AI as useful not only because it helps them produce better translations, but also because it supports how they learn during the translation process. In this study, the factor brings together items related to perceived usefulness, such as improving translation quality, vocabulary, and error identification, and items related to self-regulated learning, such as planning AI use, monitoring AI suggestions, checking outputs, reflecting, and evaluating learning. In other words, ALE represents the extent to which students experience

AI as both a practical translation aid and a scaffold for metacognitive engagement. This finding suggests that, in the HUFLIT context, usefulness and self-regulation are closely intertwined rather than clearly separated dimensions.

5. Discussion

The findings provide a coherent picture of how HUFLIT English majors experience AI-supported translation learning. Across all constructs, students reported consistently positive perceptions of AI, yet simultaneously expressed clear awareness of risks associated with overdependence and reduced independent thinking. This dual pattern runs throughout the descriptive results and is further clarified by the four-factor structure identified through EFA.

5.1. High AI Engagement and Perceived Benefits

Addressing Research Question 1, the finding that perceived usefulness and self-regulated learning converged into a single factor – labelled AI Learning Engagement (ALE), suggests that HUFLIT students do not perceive AI as a passive tool but as an active scaffold for their own metacognitive processes.

Students appear to find AI valuable precisely because it supports their learning engagement, including goal-setting, monitoring, and reflection during translation tasks. AI use among HUFLIT English majors is already widespread: mean usage frequency was 3.84 out of 5, and 92.7% of respondents had already used AI in translation courses. ChatGPT was the dominant tool (72.6%), and students primarily used AI for grammar checking (32.9%) and vocabulary search (24.2%), suggesting that AI serves mainly as a language-level support tool rather than a direct translation substitute. Perceived Usefulness recorded the highest and most consistent mean scores across the study (M = 4.15–4.16), confirming that students see AI as a genuinely beneficial resource for improving translation quality, vocabulary, error identification, and task completion.

5.2. The Composite AI Learning Engagement Factor

In relation to Research Question 3, the clean separation of Teacher AI Integration (TAI) as a distinct factor confirms that students perceive teacher guidance as a qualitatively different dimension from their own AI learning engagement.

The most theoretically significant finding is the emergence of the AI Learning Engagement (ALE) factor, which combined Perceived Usefulness and Self-Regulated Learning items into a single construct accounting for 26.373% of total variance. In the classical TAM framework, these are treated as separate dimensions; the present results suggest that HUFLIT students perceive AI as useful *because* it supports their metacognitive processes: planning, monitoring, checking, reflecting, and evaluating. Usefulness is therefore not merely about better translation outcomes but is deeply intertwined with self-regulated engagement. This finding contributes an important nuance to the TAM–SRL integration proposed in the theoretical framework and suggests that future research should apply confirmatory factor analysis (CFA) to further validate the ALE construct across different EFL and translation contexts.

5.3. Teacher AI Integration as a Distinct Factor

Also relevant to Research Question 1, the emergence of Cognitive Overreliance (CO) as a separate factor reflects students' dual awareness: they recognise AI as beneficial while simultaneously acknowledging a tendency toward over-dependence that may undermine independent thinking and critical translation decision-making.

Teacher AI Integration emerged as a clear independent factor (F2, 17.021% of variance), with the highest-rated item being "My teacher helps students understand when they should and should not rely on AI" (M = 4.33). This confirms the TPACK perspective: effective AI integration depends not only on tool availability but on teachers purposefully designing AI-supported activities, framing AI as a critical thinking scaffold, and providing explicit guidance on responsible use. The independence of TAI from the ALE factor also signals that students clearly differentiate between their own engagement with AI and the instructional practices of their teachers, confirming that teacher guidance remains a decisive factor even when AI tools are freely accessible outside class.

5.4. Cognitive Overreliance: Awareness Without Resolution

With respect to Research Question 2, the retention of Perceived Ease of Use (PEOU) as a distinct empirical factor supports the relevance of TAM in this context, confirming that accessibility perceptions constitute a meaningful and separable dimension of students' overall experience of AI-supported translation learning.

Cognitive Overreliance (CO) formed a distinct independent factor (F3, 15.407% of variance). The item "I sometimes depend too much on AI when doing translation tasks" had the highest mean in this construct (M = 4.03), and "Using AI sometimes reduces my independent thinking ability" loaded most strongly on this factor (loading = 0.880). Students are clearly aware of their overdependence tendencies; however, awareness alone does not guarantee behavioural change. The relatively high means on items about accepting AI suggestions without careful analysis suggest that cognitive overreliance is already a lived reality for many students, not merely a theoretical concern. Notably, the separation of CO from ALE implies that even students who self-regulate their AI use may simultaneously experience dependence, pointing to the coexistence of metacognitive engagement and habitual reliance in the same learner.

5.5. Perceived Ease of Use and Confidence

Taken together and addressing Research Question 2, the four-factor structure – ALE, TAI, CO, and PEOU, provides an empirically derived model of how HUFLIT students perceive AI-supported translation learning, offering a more differentiated account than prior single-framework studies.

Perceived Ease of Use (PEOU) emerged as the fourth independent factor (F4, 14.725% of variance), with means ranging from 3.93 to 4.15. Students broadly found AI tools accessible and easy to learn; however, the confidence item recorded the lowest mean (M = 3.93) and the largest standard deviation (SD = 1.014) in this group. This gap between functional ease of use and genuine AI confidence suggests that while most students can operate AI tools, a subset may not yet feel equipped to critically evaluate, selectively apply, or responsibly adapt AI output in complex translation contexts.

5.6. Theoretical Implications

This study offers three key theoretical contributions. First, the emergence of the ALE composite factor suggests that in AI-rich translation environments, perceived usefulness and self-regulated learning may converge into a single empirical factor rather than functioning as separate dimensions as classical TAM proposes. Future research should test this preliminary structure using confirmatory factor analysis (CFA) across different EFL contexts.

Second, the clean separation of Teacher AI Integration (TAI), Cognitive Overreliance (CO), and Perceived Ease of Use (PEOU) supports the value of combining TAM, TPACK, and

SRL. Each lens contributes a non-redundant dimension, technology acceptance, pedagogical practice, and metacognitive engagement, providing a more comprehensive account of AI learning than any single framework.

Third, the study contributes context-specific evidence from Vietnamese higher education, demonstrating that high generative AI adoption can coexist with a high awareness of cognitive overreliance. This underscores the need for educational technology theories to account for local pedagogical cultures and assessment norms rather than assuming universal applicability.

5.7. Practical Implications

The findings point to four practical directions for translation instruction at HUFLIT and comparable programmes.

First, given the pervasive AI use already documented, formal and intentional AI integration into translation syllabi is more productive than restriction. Tasks that require students to compare, evaluate, and justify AI output can transform AI from a shortcut into a trigger for critical reflection.

Second, the prominence of TAI highlights the need for systematic teacher development in AI literacy and TPACK. Teachers should be supported to design tasks that require critical engagement with AI, model responsible AI use, and communicate explicit guidelines on when AI is and is not appropriate.

Third, the CO factor calls for direct instructional attention to AI dependence. Process-oriented assessment tools such as translation logs, revision comparisons, and AI-use reflection notes can shift grading focus from final products to decision-making quality, rewarding critical evaluation over unreflective reliance.

Fourth, the gap between functional ease of use and AI confidence suggests students need guided AI literacy training within translation courses: how to prompt effectively, evaluate AI feedback critically, recognise errors, and disclose AI assistance responsibly. At the institutional level, clear and context-sensitive AI policies, developed in dialogue with both teachers and students, are essential to ensure that ethical AI use becomes a shared responsibility.

6. Conclusion

This study investigated HUFLIT English majors' perceptions of AI-supported EFL translation learning through an integrated TAM–TPACK–SRL framework, drawing on survey data from 219 students. The findings reveal that AI tools, particularly ChatGPT and Grammarly, have become widely and frequently adopted in translation learning, with students perceiving clear benefits in terms of translation quality, vocabulary development, error identification, and self-regulated engagement.

The EFA results produced a four-factor empirical model, with the most notable finding being the emergence of AI Learning Engagement as a composite construct that merges perceived usefulness and self-regulated learning. This suggests that in the HUFLIT context, students experience AI's pedagogical value as inseparable from their own metacognitive processes of planning, monitoring, and reflecting. Teacher AI Integration, Cognitive Overreliance, and Perceived Ease of Use emerged as three additional independent factors, confirming that students perceive teacher guidance, dependence risks, and tool accessibility as distinct and meaningful dimensions of their AI-supported learning experience.

These findings carry both theoretical and practical significance. Theoretically, the study contributes an empirically derived four-factor model and context-specific evidence for extending the TAM–TPACK–SRL framework in AI-supported translation education. Practically, the results underscore the importance of moving beyond access to AI tools toward purposeful pedagogical integration, explicit AI literacy instruction, and clear institutional policies that support responsible use.

This study is not without limitations. The sample was drawn from a single institution and was predominantly composed of second-year students, which may limit the generalizability of the findings. Future research should expand the sample across multiple universities and academic years, and apply confirmatory factor analysis and structural equation modelling to further validate and extend the empirical model developed here.

Ultimately, this study positions AI not as a threat to translation learning but as a powerful pedagogical resource whose value depends critically on how it is taught, framed, and integrated. As Vietnam continues its transition toward English as a second language, ensuring that both teachers and students engage with AI critically, responsibly, and strategically will be essential for preparing the next generation of competent and autonomous translators.

References

Alangari, T. S. (2025). The effect of AI-assisted learning on EFL writing proficiency: Quasi-experimental and cluster analysis. *Educational Process International Journal, 17*, e2025345. https://doi.org/10.22521/edupij.2025.17.345

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal, 14*(1), 3–24. https://doi.org/10.29140/jaltcall.v14n1.221

Compagnoni, I., Oguilve, V., & Wen, W. (2025). Ethical frontiers of AI in language pedagogy: Navigating transparency, accountability, and equity. *International Journal of Linguistics, 17*(4), 1. https://doi.org/10.5296/ijl.v17i4.22768

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Manzoor, S., Jamil, H., Nawaz, M., Shahbaz, M., & Farooqi, S. U. H. (2025). A systematic review of AI-integrated tools in ESL/EFL education. *Pertanika Journal of Science & Technology, 33*(5), 2125–2153. https://doi.org/10.47836/pjst.33.5.05

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Nhan, L. K., Hoa, N. T. M., & Quang, L. V. N. (2025). Leveraging AI for writing instruction in EFL classrooms: Opportunities and challenges. *Educational Process International Journal, 15*(1), e2025158. https://doi.org/10.22521/edupij.2025.15.158

Pastor, D. G. (2021). Introducing machine translation in the translation classroom: A survey on students' attitudes and perceptions. *Tradumàtica: Tecnologies de la Traducció, 19*, 47–65. https://doi.org/10.5565/rev/tradumatica.273

Rahman, W. Z. (2026). Artificial intelligence (AI) in English writing skills learning: A systematic review. *ETDC: Indonesian Journal of Research and Educational Review*, *5*(2), 981–994. https://doi.org/10.51574/ijrer.v5i2.4727

Ren, X., & Wang, R. (2025). Integrating human–AI collaboration into translation education: A comprehensive protocol for assessment, diagnosis, and strategy development. *PLOS ONE, 20*(12), e0338089. https://doi.org/10.1371/journal.pone.0338089

Saud, D. S. (2025). Generative AI in English language teaching: Opportunities and challenges. *Journal of Jayaprithvi Multiple Campus, 1*, 44–59. https://doi.org/10.3126/jjmc2.v1i1.81437

Sutrisno, A., Cahyanto, I., & Meiliasari, E. (2025). Exploring students' perceptions of Google Translate as AI-assisted translation in Indonesian vocational high school. *Wiralodra English Journal, 9*(1), 229–242. https://doi.org/10.31943/wej.v9i1.405

Tian, S., Jia, L., & Zhang, Z. (2023). Investigating students' attitudes towards translation technology: The status quo and structural relations with translation mindsets and future work self. *Frontiers in Psychology, 14*, 1122612. https://doi.org/10.3389/fpsyg.2023.1122612

Yan, Q., Hussin, S., & Jamaludin, K. A. (2025). Students' perceptions of AI-assisted blended teaching in a Chinese translation class: A preliminary study. *Arab World English Journal, CALL Special Issue* (11), 160–176. https://doi.org/10.24093/awej/call11.10

Zimmerman, B. J. (2000). Attaining self-regulation. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Bionote

Nam Anh Tran holds a Master's degree in Education and is a Lecturer in the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), Vietnam. His research interests include AI-assisted language learning and the integration of educational technology in higher education contexts. His current work examines how generative AI tools such as ChatGPT and Grammarly shape student perceptions, metacognitive engagement, and self-regulated learning in university-level EFL translation courses in Vietnam.

Thanh Tu Tran holds a Master's degree in Business Administration and is a Lecturer in the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages – Information Technology (HUFLIT), Vietnam. His research interests span artificial intelligence in education and business management, with a particular focus on how AI tools can be integrated into language learning and research in Vietnamese higher education.

s

#198 INTEGRATION OF LOCAL HISTORY AND CULTURAL MATERIAL INTO ENGLISH LANGUAGE TEACHING AS BASIS FOR THE DEVELOPMENT OF MULTICULTURAL PERSONALITY

Gulnur Polkina[1], Evgeniya Tsyganova[2], Svetlana Radionova[3]

Abstract

This study is devoted to local history and cultural material as a means of developing a multicultural identity in the foreign language classroom. The integration of such material into the subject area of "Foreign (English) Language" motivates learners to acquire English as a tool for self-expression in a multicultural world and fosters their readiness for effective intercultural communication. Understanding the cultural realities of one's own country allows learners to convey their meaning more accurately in a foreign language and to better comprehend foreign-language texts. English language learning that is connected with personal experience, history, and culture becomes more meaningful and serves as a foundation for successful intercultural interaction. Awareness of the value of one's heritage culture builds confidence in intercultural communication and contributes to the preservation of one's national identity. The study presents a positive experience in developing and piloting English language teaching content based on local history and cultural material of the Republic of Tatarstan. It also describes the teaching strategies used in English classes based on local history, culture, and traditions.

***Keywords**:* teaching a foreign language, local history and cultural material, regional culture, multicultural identity, intercultural communication

1. Introduction

Recognising the world's cultural diversity is crucial for creating an effective educational environment in foreign language teaching. Learning a foreign language aims to develop students' communicative competence, to raise awareness of the role of the foreign language as a tool for interpersonal and intercultural interaction, to promote students' overall speech development, to foster civic identity, and to broaden their outlook. In this regard, incorporating thematic units on the cultural values and traditions of one's own country into the subject area of "Foreign Language" represents a theoretically sound and methodologically relevant direction for the development of language didactics.

This aligns with the strategic goals of the Federal State Educational Standard for Basic General Education and the Concept of the State National Policy of the Russian Federation. Thematic units focused on one's heritage culture enhance motivation to learn a foreign language as a tool for effective intercultural interaction and contribute to the implementation

[1] *Ph. D., Associate Professor, Department of Romano-Germanic Languages and Methods of their Teaching, Faculty of History and Philology, Federal State Budgetary Educational Institution of Higher Education "Naberezhnye Chelny State Pedagogical University", Naberezhnye Chelny, Russia, e-mail: gulnurpolkina@gmail.com*

[2] *Ph. D., Associate Professor, Department of Romano-Germanic Languages and Methods of their Teaching, Faculty of History and Philology, Federal State Budgetary Educational Institution of Higher Education "Naberezhnye Chelny State Pedagogical University", Naberezhnye Chelny, Russia*

[3] *Ph.D., Associate Professor, Department of Romano-Germanic Languages and Methods of their Teaching, Faculty of History and Philology, Federal State Budgetary Educational Institution of Higher Education "Naberezhnye Chelny State Pedagogical University", Naberezhnye Chelny, Russia*

of state policy aimed at preserving civic and national identity in a changing world. Incorporating culturally familiar content into the school subject of "Foreign Language" encourages students to reflect on similarities and differences between the culture of the target language country and their own country or region, which in turn fosters a deeper understanding of their own cultural and behavioural characteristics as well as those of "others."

Thus, the relevance of this study is driven by the need to find effective methods for integrating regional cultural content into the English language teaching process in lower secondary school—methods that not only increase students' motivation but also prepare them for intercultural communication without losing their own cultural identity.

2. Literature Review

2.1. Foundational Theories of Intercultural and Communicative Language Teaching

The current model of foreign language teaching, based on the principles of the communicative approach (Hymes, 1972; Canale & Swain, 1980) and intercultural communication (Byram, 1997), requires a shift from a monocultural to a multicultural model of instruction. Teaching foreign languages in a globalised world, as noted by Western and Russian methodologists (Kramsch, 1993; Byram, 2002; Myasnikova, 2013; Tarnaeva, 2015), demands an integrated approach that includes a cultural component, digital technology, and consideration of learners' personal experiences related to their background.

Kramsch (1998) highlighted the importance of the sociocultural context in education, including foreign language teaching. She argued for the necessity of developing learners' capacity for critical cultural reflection – understanding one's own culture through the lens of the foreign language being studied. Kramsch maintained that in foreign language teaching, it is important to develop learners' ability to analyse and interpret "other" culture rather than merely absorbing its stereotypes. Such an approach enables learners to navigate various intercultural situations more effectively and to engage in intercultural dialogue.

2.2. Regional Cultural Content and Identity Formation in Foreign Language Teaching

Studying one's heritage culture in a foreign language equips learners with the ability to present their country and region in a global context, which aligns with the principles of the constructivist theory of learning proposed by the renowned Soviet scholar L. S. Vygotsky (Vygotsky, 1978). Norton (2013) argued that learning a foreign language influences the formation of imagined communities. Risager (2007) addressed the "principle of cultural symmetry" and proposed a model of bilateral cultural exchange, whereby learning about a foreign culture is mediated through reflection on one's own culture.

Russian researchers in cultural linguistics (Vorobyeva, 2003; Maslova, 2001; Tarnaeva, 2015; Mushtavinskaya, 2018; Silonyan, 2025) argued that studying one's heritage culture through a foreign language is a way of actualising language as a sign of culture. Linguistic material related to one's "own" culture is acquired by students more consciously and with greater emotional significance because it stems from their natural environment – what they already know well. Studying the values, traditions, and history of one's people in a foreign language gives learners the opportunity to present their country internationally and to overcome language barriers when explaining their homeland's cultural realities. Thus, English serves as a tool for intercultural communication rather than one of "assimilation," thereby strengthening national identity.

In recent years, international research (Wang et al., 2022; Haerazi et al., 2023; Yuliantari et al., 2023; Azhary et al., 2024; Reformana et al., 2024; Patanduk et al., 2025) has

demonstrated sustained interest in the problem of integrating students' cultural backgrounds into language education. Integrating cultural values supports character development, classroom inclusivity, and meaningful learning. Incorporating local cultures into English teaching materials is crucial for promoting cultural diversity, cultural sensitivity, and enhancing learners' language learning experience. By including local elements, teachers can make learning content more relatable and engaging for students, ultimately leading to a more inclusive and effective educational environment.

3. Methodology

Content Selection. The selection of linguocultural material for Grade 8 was based on two complementary principles. The principle of regional identity required in-depth presentation of local cultural facts (history, geography, ethnography of the native region). The principle of multicultural orientation aimed at comparative analysis with cultures of different countries. To resolve the tension between these principles, only those facts from local culture that have comparable phenomena in other cultures (holidays, national cuisine, urban spaces, famous people) were selected. This made it possible to use local history material as a starting point for multicultural comparison rather than as an alternative to it.

Development of Teaching Materials. To address the dilemma of "accessibility versus authenticity" for learners aged 14 to 15 years, a combined approach was used. The introduction of a cultural concept was carried out through a teacher-generated semi-authentic text. This was followed by exposure to limited authentic fragments: captions to illustrations, short facts from encyclopaedias, QR codes linking to authentic guidebooks. This combination avoided oversimplifying cultural content (which leads to stereotyping "tradition" as merely holidays and food) while simultaneously preventing cognitive overload caused by overly complex lexical and syntactic structures.

Research Design. This study follows a pilot (work-in-progress) design and represents a descriptive-empirical work with elements of quasi-experimental observation. The main objective was to test the developed teaching modules and collect preliminary data for a subsequent full-scale experimental study. The research did not include a control group or random assignment of participants.

Participants. The study involved 28 Grade 8 students (aged 14–15 years) from Secondary School No. 41 in Naberezhnye Chelny (Republic of Tatarstan, Russia). The class had a language-focused curriculum, which entailed advanced study of English (five hours per week instead of three hours in the basic curriculum). The participants' English proficiency level corresponded to A2 (Pre-Intermediate) on the CEFR scale. All participants were bilingual (fluent in Tatar and Russian), which is a characteristic feature of the region. Informed consent was obtained from the parents or legal guardians of all participants; participation was voluntary, and anonymity of results was guaranteed.

Context and Setting. The study was conducted within extracurricular activities (elective English classes) from September to December 2025. The total intervention lasted 16 weeks (one semester). The extracurricular format was chosen to avoid conflict with the compulsory curriculum (the standardised "Spotlight-8" teaching package) and to allow flexible integration of the developed local history modules.

Instruments and Materials. The study used the following materials, developed by the research project group "Tatarstan: Language and Culture, History and Modernity" at the Department of Romance-Germanic Languages and Teaching Methods, Naberezhnye Chelny State Pedagogical University: three thematic modules – "Welcome to Tatarstan," "National

Holidays of Tatarstan," and "Where the History of Tatarstan Begins." Within each module, one integrated lesson was conducted. Tasks included non-linear texts (infographics, diagrams, timelines), video materials (reports on Tatar diasporas abroad), and coded texts (with reversed letters, redaction, picture-based encryption). After completing the three modules, an intellectual linguocultural quiz called the "Tatar-lish Quiz" (five categories with four questions of varying difficulty levels) was administered as a final assessment tool.

Data Collection Procedure. Data collection was carried out in three stages. At the preliminary stage (Week 1), an initial questionnaire on attitudes towards learning English through local culture was administered (qualitative data not included in this article). At the main stage (14 weeks), integrated lessons based on the module content were developed and delivered within extracurricular activities, with field notes, photo and video materials, and timing of coded text completion recorded. At the final stage (Week 16), the Tatar-lish Quiz was conducted.

Data Analysis. Quantitative analysis included calculating the percentage of correct answers for each of the five quiz categories (N = 28) and calculating the average decoding time per sentence (in seconds) for the first and fifth coded text tasks. Comparison of means was performed without statistical testing due to the pilot nature of the study. Qualitative analysis included the analysis of field notes (students' reactions, communication barriers, engagement) and the analysis of oral reflections after the lessons.

Methodological Limitations. Due to the pilot nature of the study, the methodology has the following limitations: (a) absence of a control group, (b) small sample size (N = 28), (c) lack of randomisation, and (d) possible "special interest" effect (linguistically inclined students may be more motivated from the outset). All limitations will be addressed when planning the main experimental study.

4. Results

In intercultural communication, the sociocultural component of communicative competence plays a significant role in the formation of a plurilingual personality. This component involves developing the ability to present one's heritage culture in a foreign language and to compare the historical and cultural features of one's region of residence with those of the target language countries. The Federal State Educational Standard for Basic General Education "ensures the preservation and development of cultural diversity and linguistic heritage of the multinational people of the Russian Federation."

Within the developed modules "Welcome to Tatarstan," "National Holidays of Tatarstan," and "Where the History of Tatarstan Begins," the following types of tasks were used: inquiry-based tasks, group research work, and project-based activities. For example:

1. Hold a survey among your classmates to find out what nationality they belong to. Present the results in graphic form.

2. Project: Work in groups. Create a seasonal calendar of Tatar holidays to present to the class.

3. Web quest: Search the internet for additional information about the history of the Kazan Kremlin. Draw a chronological line of its history, noting only the most important events.

The Tatar-lish Quiz was piloted with Grade 8 students with a language-focused curriculum (N = 28). The quiz consisted of five categories: "Culture and Traditions," "Folklore," "Famous People," "Linguistic Features," and "Geography." Each category

included four questions of varying difficulty levels. The quiz format made it possible to assess knowledge of both English and Tatar languages.

Table 1

Quiz Results of Grade 8 Students (N = 28)

Category	Success Rate
Culture and Traditions	77%
Linguistic Features	60%
Folklore	55%
Geography	53%
Famous People	32%

Students demonstrated high results in the categories “Culture and Traditions” (77% correct answers) and “Linguistic Features” (60% correct answers). At the same time, areas of academic deficit were identified: students showed low levels of knowledge in the categories “Folklore,” “Geography,” and “Famous People” (55%, 53%, and 32% correct answers, respectively). The notably low result in the “Famous People” category (32%) warrants particular attention, as it indicates insufficient integration of information about regional figures in culture, science, and sports into the educational process.

Integrated English and Tatar language lessons were conducted within extracurricular activities. Students were given tasks involving non-linear texts, video materials, and coded texts. The results of task completion showed that non-linear texts develop predictive skills and the ability to condense information semantically. Video materials contributed to the formation of listening and pronunciation skills, situational motivation, and visual fixation of culturally marked vocabulary. Coded texts (with reversed letters, redaction, picture-based encryption) activated mechanisms of linguistic guesswork and anticipation: the average decoding time decreased from 90 seconds (first task) to 45 seconds (fifth task), indicating the formation of probabilistic forecasting strategies.

In accordance with the principle of problem-based learning (developing critical thinking by presenting ambiguous cultural facts), tasks included elements comparing different points of view. Preliminary observations showed that excessive problematisation at the initial stage (for example, discussing controversial pages of regional history) created a communication barrier. Therefore, only those problem situations that allowed for different interpretations within a value-neutral field – so-called “consensus-discussion” topics (e.g., comparing Tatar and British harvest festival traditions) – were used in the main part of the lessons. Students successfully completed such tasks, demonstrating readiness to express and argue their opinions in English.

5. Discussion

The proposed approach to developing learning content proceeds from the premise that local history and cultural studies material acts not as a peripheral addition to the “main” foreign language content but as a system-forming factor. It is through making sense of facts of one’s heritage culture in a foreign language that the transition occurs from formal acquisition of vocabulary and grammar to personally meaningful expression.

Analysis of Quiz Results. The results presented in Table 1 revealed differentiation in the level of mastery of various categories of cultural content. The high level in the "Culture and Traditions" category (77%) can be explained by the fact that this material is most frequently presented in standard teaching packages (holidays, food, customs) and is also actively disseminated in the media. The low level in the "Famous People" category (32%) requires special attention: this may be due to the fact that English language textbooks predominantly present information about figures from English-speaking countries rather than regional celebrities.

This gap aligns with the conclusion of the systematic review by Wang et al. (2022) that globally distributed teaching materials often neglect local cultural elements, creating an imbalance in favour of the target language countries. Non-systematic inclusion of local history material in the educational process is the cause of an information gap that does not provide the necessary effect of culture-oriented teaching.

The obtained results confirm and complement the findings of international research. As in the study by Patanduk et al. (2025), we observe that students demonstrate higher engagement and success in those categories where local content is closely related to their everyday experience (traditions, language). The study by Haerazi et al. (2023) correlates with our findings that integrating local history material into English language teaching enables learners to navigate intercultural communication with greater proficiency and cultural sensitivity. This approach contributes to the ongoing advancement of language education and its transformative impact on learners' communicative capabilities and cultural understanding.

One of the key conclusions of Reformana et al. (2024) finds empirical confirmation in our context. The use of local cultural topics allows learners to connect with their own heritage, which not only enhances their motivation to participate in speaking activities but also boosts their confidence in expressing ideas in English. This familiarity with the content encourages a natural flow of conversation, reducing the cognitive load often associated with learning foreign language structures and vocabulary.

Resolving Methodological Contradictions. The results of the study allow us to propose specific ways of resolving the three contradictions stated in the methodology.

First contradiction (regional identity vs. multicultural orientation). This is resolved through the principle of "translatability" of cultural facts, implemented in the quiz: each category contained questions about both Tatarstan and English-speaking countries. However, the low results in the "Folklore" and "Geography" categories (55% and 53%) indicate that this is insufficient. A more systematic approach is needed – selecting not any comparable phenomena but those that have high frequency in both cultural contexts (for example, comparing the Kazan Kremlin and the Tower of London as symbols of statehood rather than merely "sights").

Second contradiction (standardised textbooks vs. regional specificity). This is resolved through the proposed mixed model: an introductory module (two to three hours) "Talking about our small homeland in English" → distributed tasks (one to two per each textbook topic) → a final project at the end of the year. The results of the conducted lessons show that distributed tasks should not be mere "insertions" into textbook topics but specially designed mini-modules accompanied by QR codes linking to authentic regional resources (as was done with the coded texts). Students positively evaluated this format, confirming the effectiveness of the approach.

Third contradiction (age-related abilities vs. material complexity). This is resolved through a combined approach to texts. The data show that semi-authentic texts (teacher-generated) are effective for introducing cultural concepts, while short authentic fragments are effective for developing linguistic guesswork. At the same time, the average decoding time for coded texts decreased from 90 to 45 seconds, indicating rapid formation of probabilistic forecasting strategies.

Age Characteristics and the Activity-Based Model. Contemporary research confirms the findings of Vygotsky (1978) and Elkonin (1989) on the significance of adolescence for the formation of reflection and identity (Shapovalenko et al., 2026). A Grade 8 student is already capable of abstract thinking and comparisons but remains sensitive to the evaluation of their personal position. This creates both favourable conditions for including local history material (since the topic of "my small homeland" is personally significant) and risks: poorly chosen material or imposing a particular view of history or culture may be rejected as "officialdom" or propaganda. Consequently, learning content should be structured not as a transmission of predetermined "correct" cultural facts (an information model) but as a space for expression and discussion (an activity-based model). The obtained results confirm this thesis: students actively engaged in discussing those cultural topics where expressing their own position was allowed, and conversely demonstrated passivity when an attempt was made to impose a single correct view.

Ethnocultural Relevance. The problem of ethnocultural relevance of learning content requires special attention: in classes with a multi-ethnic composition, local history material from one group of students may be alien to another. The solution is to include material related to the regional space common to all, rather than to the ethnic traditions of one group. When addressing ethnocultural topics, a format of intergroup comparison should be used, where each group presents its own material. In the Republic of Tatarstan, where two state languages (Russian and Tatar) function, this principle takes on particular significance: students learn English by comparing it with their native (Tatar) and Russian languages, which contributes to the development of a plurilingual personality.

6. Conclusion

Including familiar linguocultural content in the school subject of "Foreign Language" encourages students to reflect on similarities and differences between cultures, contributing to a deeper understanding of their own cultural characteristics and those of "other" cultures. The integration of local history material into Grade 8 English language teaching, using the example of the Republic of Tatarstan, demonstrates increased motivation, the formation of a plurilingual identity, and the strengthening of national identity. Students reported forming a stronger emotional connection with the material, perceiving English not as an external imposition but as a medium that reflects their own lives.

Further work within the project involves the creation of a teaching guide for English teachers containing up-to-date local history material for Grades 7 and 8 of general education schools in the Republic of Tatarstan and an empirical verification of the effectiveness of the proposed mixed model. The next phase of the study will include a quasi-experiment with a larger sample (N = 120), a control group, and a pre-test/post-test design, which will yield statistically significant data on the effectiveness of the developed modules and tasks.

Acknowledgments

The authors would like to thank the administration and teachers of Secondary School No. 41, Naberezhnye Chelny, for their support in conducting this study. Special gratitude is

extended to the research project group “Tatarstan: Language and Culture, History and Modernity” at Naberezhnye Chelny State Pedagogical University for developing the teaching modules. The authors also express their appreciation to the students and their parents for their voluntary participation and trust.

References

Azhary, L. (2024). Integrating local cultures in English teaching materials. Journal of Language and Cultural Education, 12(3), 2040–2050.

Byram, M. (1997). Teaching and assessing intercultural communicative competence. Multilingual Matters.

Byram, M. (2002). Foreign language education as political and moral education: The European idea. In L. H. (Ed.), Learning for democracy, 105–116. UCL Press.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1(1), 1–47.

Elkonin, D. B. (1989). Izbrannye psikhologicheskie trudy [Selected psychological works]. Pedagogika.

Haerazi, H., Irawan, L. A., & Sotlikova, R. (2023). Interculture-based language learning model to improve prospective English teachers' speaking skills viewed from linguistic awareness. JEELS (Journal of English Education and Linguistics Studies), 10(2), 471–497. https://doi.org/10.30762/jeels.v10i2.1763

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), Sociolinguistics: Selected readings, 269–293. Penguin.

Kramsch, C. (1993). Context and culture in language teaching. Oxford University Press.

Kramsch, C. (1998). Language and culture. Oxford University Press.

Maslova, V. A. (2001). Lingvokul'turologiya [Linguoculturology]. Academia.

Mushtavinskaya, I. V. (2018). Tekhnologiya razvitiya kriticheskogo myshleniya na uroke i v sisteme podgotovki uchitelya [Technology of critical thinking development in the classroom and in teacher training]. KARO.

Myasnikova, O. V. (2013). Formirovanie mezhkul'turnoy kompetentsii pri obuchenii inostrannomu yazyku v vuze [Formation of intercultural competence in teaching a foreign language at university]. Vestnik OGU, 2(163), 45–49.

Norton, B. (2013). Identity and language learning: Extending the conversation (2nd ed.). Multilingual Matters.

Patanduk, S., Dollah, S., Sunra, L., Rachel, R., Sirande, N., & Duapadang, A. (2025). Incorporating local cultural values in English teaching materials: A lecturers' perception. Jurnal Studi Guru dan Pembelajaran, 8(2), 1112–1121. https://doi.org/10.30605/jsgp.8.2.2025.6159

Reformana, G., Haerazi, H., & Irawan, L. A. (2024). Sasaq local wisdom-based ELT materials in improving intercultural competences and speaking skills as Pancasila learners' profiles. Journal of English Education and Teaching, 8(4), 764–782. https://doi.org/10.33369/jeet.8.4.764-782

Risager, K. (2007). Language and culture pedagogy: From a national to a transnational paradigm. Multilingual Matters.

Shapovalenko, I. V., Kochetova, Yu. A., Sem'ya, G. V., Merkul', I. A., & Radchikov, A. S. (2026). Lichnostnaya identichnost' sovremennykh podrostkov: rezul'taty obshcherossiyskogo issledovaniya [Personal identity of modern adolescents: Results of

an all-Russian study]. Psikhologicheskaya nauka i obrazovanie, 31(1), 141–157. https://doi.org/10.17759/pse.2026000005

Silonyan, A. A. (2025). Development of cultural identity of schoolchildren in foreign language lessons in basic comprehensive school. World of Science. Pedagogy and Psychology, 13(1), 61PDMN125.

Tarnaeva, L. P. (2015). Obuchenie inostrannym yazykam v kontekste dialoga kul'tur [Teaching foreign languages in the context of dialogue of cultures]. Izvestiya RGPU im. A.I. Gertsena, 175, 112–117.

Vorobyeva, I. V. (2003). Lingvokul'turologicheskiy podkhod k obucheniyu inostrannym yazykam [Linguoculturological approach to teaching foreign languages]. RIO GAGU.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wang, Z., & Wang, H. (2022). Culturally-oriented training of highly qualified specialists in Russian: Concept and practice. Philological Class, 27(2), 141–148. https://doi.org/10.51762/1FK-2022-27-02-14

Yuliantari, S., & Huda, T. (2023). Integration of culturally-responsive teaching in English learning. Pubmedia Jurnal Pendidikan Bahasa Inggris, 1(1), 8. https://doi.org/10.47134/jpbi.v1i1.17

Bionote

Gulnur Polkina is an Associate Professor at the Department of Romano-Germanic Languages and Methods of their Teaching, the dean of the Faculty of History and Philology in Naberezhnye Chelny State Pedagogical University, Russia. Her research interests include intercultural communication, identity formation in foreign language education, and integrating regional cultural content into ELT.

Evgeniya Tsyganova is an Associate Professor, the Head of the Department of Romano-Germanic Languages and Methods of their Teaching. Her work focuses on bilingual education, linguoculturology, and polylingual competence development of students.

Svetlana Radionova is an Associate Professor of the Department of Romano-Germanic Languages and Methods of their Teaching. She specialises in communicative language teaching, implementing interactive and student-centered teaching methods.

#201 INTEGRATING DEBATE-BASED PEDAGOGY TO SUPPORT VIETNAM'S TRANSITION FROM EFL TO ESL

Nguyen Thuy Hong Van[1], Guillermo Reyna[2]

Abstract

Vietnam is undergoing a major change in how English is taught. People in charge of education are moving away from the old way of teaching English as a foreign language. Instead, they want to teach English as if it were a second language people use every day. This new way of teaching focuses on using English to communicate, rather than just learning grammar rules. But even with this change, many classrooms still focus too much on grammar and exam preparation. This means students don't get to practice using English in real-life situations or think critically about what they're learning. This study examines how using debates in the classroom can help address this problem. It can help students improve their English communication and think more deeply about what they're learning. This research examines how classroom debates can help students express their opinions, examine evidence, and respond to others' views in English. It also checks if debating in class can improve critical thinking and make the classroom a better place to learn English. The study draws on ideas from teaching argumentation and language skills to examine how class debates affect students' participation and engagement. By doing so, it aims to determine whether debating can help Vietnamese students improve their English-speaking and critical-thinking skills. The goal is to create a classroom where students feel comfortable using English and can develop their skills in a fun and interactive way. This research aims to show that using debates in teaching can help students feel more confident when communicating, think more critically, and use language more effectively. The study provides teachers and curriculum designers with new ideas for teaching English in ways that align with Vietnam's evolving goals for English language education. The findings also show that successful debate implementation depends on structured scaffolding, flexible classroom formats, and teacher support. While teachers strongly valued debate, many used it only occasionally, indicating a belief-practice gap stemming from limited time, large class sizes, exam-oriented curricula, and insufficient practical training.

***Keywords*:** debate, critical thinking, argumentative, EFL, ESL, communication, teacher training.

1. Introduction

Vietnam is currently experiencing a significant shift in English language education. Traditionally, English has been taught largely as a foreign language, with classroom practices often emphasising grammar, vocabulary memorisation, examination preparation, and teacher-led instruction. However, recent policy directions show a stronger national aspiration to move English education beyond conventional foreign-language instruction. Decision No. 2080/QĐ-TTg continued Vietnam's national foreign language education reform for the period 2017–2025, while Decision No. 2371/QĐ-TTg approved the project of "making English the second language in schools" for the period 2025–2035, with a vision to 2045 (Prime Minister of Vietnam, 2017, 2025). This policy direction reflects the growing expectation that English

[1]*Corresponding author, Dr., INTESOL Vietnam - Veducon, Ho Chi Minh City, Vietnam; Email: van.nguyenthuyhong75@gmail.com*
[2] *INTESOL Vietnam - Veducon, Ho Chi Minh City, Vietnam*

should be used more widely and effectively in teaching, communication, management, and school activities, creating an English-speaking ecosystem across educational levels.

The transition from English as a Foreign Language (EFL) toward English as a Second Language (ESL) should not be understood simply as an increase in the number of English lessons. More importantly, it involves a change in English's role in education. English is expected to move from being mainly a school subject studied for tests to becoming a language that learners use for communication, learning, reasoning, and participation. This orientation is consistent with communicative approaches to language education, which emphasise the ability to use language appropriately and effectively in social contexts, rather than merely mastering grammatical forms (Hymes, 1972; Canale & Swain, 1980). Hymes' concept of communicative competence challenged the view that language competence is only grammatical knowledge, while Canale and Swain's model expanded communicative competence to include grammatical, sociolinguistic, discourse, and strategic competence.

In this context, Vietnam's EFL-to-ESL transition requires pedagogical approaches that create more meaningful opportunities for students to use English as a functional tool. Approaches such as Content and Language Integrated Learning (CLIL) and English as a Medium of Instruction (EMI) have often been discussed as possible pathways because they position English not only as the object of study but also as a medium for learning content. CLIL, for example, is commonly understood as a dual-focused educational approach in which an additional language is used to teach and learn both content and language (Coyle, Hood, & Marsh, 2010). Similarly, EMI reflects the use of English to teach academic subjects in contexts where English is not necessarily the first language of students. These approaches therefore represent a broader educational movement in which English functions as a medium for learning, thinking, and communication.

However, while CLIL and EMI provide valuable institutional and curricular models, they may not always be easy to implement widely in regular English classrooms. Successful implementation often requires teacher preparation, appropriate materials, institutional support, and sufficient learner readiness (Thi Thu Hien, 2022). Therefore, there is a need to explore classroom-level pedagogical strategies that can support the same underlying goal: helping learners use English meaningfully, not only study it formally. This paper argues that debate-based pedagogy offers one such practical and accessible pathway. Like CLIL and EMI, debate shifts the role of English from a subject of study to a medium of use. However, rather than using English primarily to learn subject content, debate enables students to use it as a tool for expressing viewpoints, constructing arguments, evaluating evidence, responding to opposing ideas, asking questions, and engaging in critical discussion.

Debate-based pedagogy is particularly relevant to English language classrooms because it combines communicative language use with cognitively demanding learning. Previous studies have shown that debate can support speaking development, critical thinking, argumentation, and student engagement in EFL/ESL contexts (Aclan & Aziz, 2015; Iman, 2017; El Majidi et al., 2024). Debate activities require learners to speak with a clear communicative purpose, organise ideas logically, listen to others, respond spontaneously, and use persuasive language. As El Majidi et al. argue, debate lends itself to student-centred pedagogy because students do much of the talking and thinking during the learning process. In this sense, debate is not merely a speaking activity or a competitive event; it is a structured pedagogical practice that can help students use English for meaningful interaction and higher-order thinking.

This is especially important in Vietnamese English classrooms, where many students may have studied English for years but still have limited opportunities to express opinions, defend viewpoints, respond to disagreement, or participate confidently in academic discussions. If Vietnam aims to make English function more like a second language in schools, learners need classroom environments where English is used not only for exercises, but also for authentic communication, reasoning, and participation. Debate can contribute to this goal by creating English-rich classroom conditions in which students use English to think, interact, and communicate with a clear purpose.

Therefore, this study examines how debate-based pedagogy can support Vietnam's transition from EFL to ESL by enhancing students' communicative competence and critical thinking in Vietnamese English language classrooms. Specifically, it explores how debate activities help students express opinions, construct arguments, use evidence, respond to opposing views, and participate more actively in classroom interaction. The central argument of this paper is that debate-based pedagogy can contribute to Vietnam's EFL-to-ESL transition by transforming English from a grammar-focused school subject into a communicative and cognitive tool. While CLIL and EMI use English as a medium for learning subject content, debate-based pedagogy uses English as a medium for argumentation, inquiry, and critical communication.

2. Literature Review

2.1 Communicative competence in TESOL

The concept of communicative competence is central to TESOL because it shifts the goal of language learning from knowing about language to using it appropriately and effectively in meaningful contexts. Hymes (1972) introduced communicative competence as a response to approaches that treated language competence primarily as grammatical knowledge. For Hymes, language users need to know not only whether an utterance is grammatically possible, but also whether it is appropriate and effective in a particular social situation. This view provides an important foundation for communicative language teaching because it emphasises language use, social meaning, and contextual appropriateness.

Canale and Swain (1980) further developed the concept by proposing a model of communicative competence that includes grammatical, sociolinguistic, and strategic competence. Canale (1983) later elaborated the model by identifying discourse competence as another important component. This broader understanding is useful for the present study because debate requires learners to do much more than produce grammatically accurate sentences. In debate, students must organise ideas coherently, respond to others, use appropriate language for agreement and disagreement, and manage communication when they lack vocabulary or need time to think.

A later pedagogically oriented model by Celce-Murcia et al. (1995) also emphasises that communicative competence involves interconnected dimensions, including discourse, linguistic, actional, sociocultural, and strategic competence. These dimensions are highly relevant to classroom-based debate. Debate requires learners to use linguistic resources accurately, organise extended spoken discourse, and perform communicative actions such as persuading and challenging, communicate appropriately with others, and use strategies to maintain interaction when communication becomes difficult.

In English language classrooms, communicative competence therefore involves several interrelated dimensions (Celce-Murcia et al., 1995). Grammatical competence helps learners express ideas accurately. Discourse competence enables them to connect ideas logically

through sequencing, explanation, contrast, and conclusion. Sociolinguistic competence helps students choose appropriate language for respectful disagreement, clarification, and persuasion. Interactional competence allows learners to listen, respond, ask questions, and negotiate meaning with others. Strategic competence helps learners continue communication when they face difficulties, such as hesitation, lack of vocabulary, or misunderstanding.

This broader understanding of communicative competence is particularly important in the Vietnamese context, where many learners may study English for years but still have limited opportunities to use it for extended interaction, argumentation, or academic discussion. If Vietnam aims to move toward an ESL-oriented learning environment, students need classroom activities that help them use English not only accurately, but also purposefully, interactively, and strategically. Debate-based pedagogy is relevant because it creates communicative situations in which learners must use English to express positions, justify claims, respond to counterarguments, and interact with peers.

2.2 English as a tool for learning, thinking, and communication

The shift from EFL to ESL-oriented education is closely connected to the changing function of English in the classroom. In a traditional EFL model, English is often treated as a subject to be studied, tested, and corrected. In a more ESL-oriented approach, English becomes a tool for communication, learning, thinking, and participation. This distinction is important because the aim is not simply to increase the number of English lessons, but to provide meaningful opportunities for learners to use English in their educational lives.

Approaches such as Content and Language Integrated Learning (CLIL) and English as a Medium of Instruction (EMI) reflect this broader direction. CLIL is commonly defined as a dual-focused educational approach in which an additional language is used to teach and learn both content and language (Coyle, Hood, & Marsh, 2010). In CLIL, learners use the target language not only to learn vocabulary and grammar, but also to understand, discuss, and process subject content. Coyle et al. (2010) emphasise that CLIL involves learning content with and through an additional language, which positions language as a medium for cognitive and academic development.

EMI is another example of this shift. Dearden (2014) defines English Medium Instruction as the use of English to teach academic subjects in countries or jurisdictions where the first language of the majority of the population is not English. EMI therefore reflects the increasing role of English in education, internationalisation, and access to academic knowledge. However, EMI also requires institutional support, teacher readiness, learner preparedness, and appropriate materials, which may make it difficult to implement fully in all educational contexts.

Although debate-based pedagogy is not the same as CLIL or EMI, it shares an important principle with them: English is used as a medium rather than only as an object of study. While CLIL and EMI use English as a medium for learning subject content (Coyle et al., 2010; Dearden, 2014), debate uses English as a medium for argumentation, inquiry, and reasoning (Kuhn & Udell, 2003) and for critical communication (El Majidi et al., 2024). This connection is important for the present study because debate offers a classroom-level pathway for creating meaningful English use even in contexts where full CLIL or EMI implementation may not yet be possible.

From a second language acquisition perspective, meaningful language use is also important because learners develop language through interaction and output. Long's (1996) interaction hypothesis emphasises the role of interaction, negotiation of meaning, and feedback

in second language development. Swain's (2005) output hypothesis highlights the importance of language production, arguing that learners need opportunities to produce language because output can push them to process language more deeply. In debate activities, students are pushed to produce extended spoken output, clarify meaning, respond to others, and adjust their language to communicate effectively. These processes make debate a relevant pedagogy for developing an English-rich classroom environment.

2.3 Debate-based pedagogy in English language classrooms

Debate-based pedagogy can be understood as the structured use of debate activities to support learning. In English-language classrooms, debate is not limited to formal competitions; it can include mini-debates, pair debates, group debates, argument-building tasks, rebuttal practice, opinion-line activities, and structured classroom discussions. What distinguishes debate from ordinary speaking practice is that students must take a position, justify it with reasons and evidence, listen to opposing views, and respond critically.

In this study, 'mini-debate' refers to short, scaffolded classroom tasks that preserve the core logic of debate—position, reason, evidence, response, and reflection—without requiring the full structure of formal competitive debate. This approach is consistent with the movement toward integrating debate-based pedagogy as an inclusive, effective tool for classroom-based critical thinking development (Akerman & Neale, 2011). These formats are especially relevant in EFL classrooms because they allow teachers to integrate argumentation into ordinary lessons without needing extended preparation time or competition-style procedures.

Research has increasingly examined debate as a pedagogical tool in EFL and ESL classrooms. Aclan and Aziz (2015) explored parliamentary debate as a pedagogical tool for developing English communication skills in EFL/ESL classrooms. Their study is relevant because it demonstrates that debate can be structured to encompass pre-debate preparation, the debate itself, and post-debate reflection. This staged process is important for language classrooms because it allows teachers to scaffold students' language, content knowledge, and interactional skills before asking them to debate.

Aclan et al. (2016) also examined debate as a pedagogical tool to develop soft skills in EFL/ESL classrooms. Their work describes how debate can support skills such as communication, teamwork, critical thinking, and confidence. For the present study, this is significant because debate is not treated only as a competitive activity, but as a classroom pedagogy that can support both language development and broader learner development.

Debate is especially suitable for communicative language teaching because it gives students a clear reason to speak. Unlike some classroom speaking tasks that may focus mainly on controlled practice or fluency drills, debate requires learners to use language for a genuine communicative purpose: to persuade, question, challenge, explain, and respond. El Majidi et al. (2024) found that debate can serve as a pedagogical tool for developing speaking skills in second language education because it scaffolds speaking processes and requires learners to use both prepared and more spontaneous speech.

In this sense, debate is both communicative and cognitively demanding. Students must organise their ideas, select appropriate expressions, provide reasons, support claims, and manage interaction with others. This makes debate a student-centred activity because learners do much of the thinking, speaking, listening, and responding. The teacher's role shifts from knowledge transmitter to facilitator, language supporter, and discussion manager. This is particularly relevant to Vietnam's EFL-to-ESL transition because the classroom needs to

become a place where students actively use English, rather than only receive explanations about English.

2.4 Debate and critical thinking

Critical thinking is closely related to debate because debate requires learners to examine ideas rather than simply express opinions. In a debate, students must make claims, support them with reasons and evidence, evaluate opposing arguments, identify weaknesses, and construct rebuttals. These processes correspond to higher-order thinking skills such as analysis, evaluation, justification, and perspective-taking.

Research on argumentation supports the link between debate and critical thinking. Kuhn and Udell (2003) examined the development of argument skills and found that argumentation can be developed through instructional intervention. Their work is relevant because debate provides repeated opportunities for students to practise constructing arguments, presenting counterarguments, and engaging in evidence-based reasoning. Rather than seeing critical thinking as an abstract skill, debate makes it visible through students' spoken reasoning, questioning, use of evidence, and rebuttal.

Iman's (2017) study is particularly relevant to EFL contexts. Using a quasi-experimental design, Iman investigated debate instruction in an EFL classroom and found that debate significantly improved students' critical thinking and speaking skills. This supports the argument that debate can contribute to both language development and cognitive development. For the present study, this is important because communicative competence and critical thinking are not treated as separate outcomes. Instead, debate is viewed as a pedagogy through which students use English to think critically and communicate more effectively.

In English-language classrooms, this connection between debate and critical thinking has an additional benefit: students practise critical thinking in the target language. They do not only learn vocabulary related to argumentation; they use English to perform cognitive actions such as explaining, comparing, justifying, questioning, challenging, and concluding. This supports the broader aim of using English as a tool for thinking and communication.

2.5 Research gap

The literature suggests that communicative competence involves much more than grammatical accuracy, and that meaningful language use is essential for developing learners' ability to communicate in real contexts. CLIL and EMI show how English can function as a medium for learning, while debate-based pedagogy offers a practical classroom-level strategy for using English as a medium for argumentation, inquiry, and critical communication. Existing studies also indicate that debate can support speaking development, interaction, argumentation, and critical thinking in EFL/ESL contexts.

However, more research is needed into how debate-based pedagogy can support Vietnam's transition from EFL-oriented instruction to an ESL-oriented learning environment. In particular, limited attention has been given to how debate may simultaneously enhance communicative competence, critical thinking, and English-rich classroom interaction among Vietnamese learners. This study addresses that gap by examining debate not merely as a speaking activity or competition format, but as a communicative and cognitively demanding pedagogy that can help students use English meaningfully for reasoning, interaction, and participation.

3. Methodology

3.1 Research Design

This study adopts a classroom-based action research design supported by mixed methods. Action research is appropriate for this study because it investigates a pedagogical intervention implemented in a real classroom context, with the aim of improving teaching practice and understanding of learner development. Action research commonly involves cycles of planning, action, observation, and reflection, allowing teachers and researchers to examine how classroom practices can be improved through systematic inquiry (Kemmis et al., 2014).

The use of a mixed-methods approach is also appropriate because the study examines both measurable and interpretive aspects of learning. Quantitative data, such as questionnaire responses and rubric scores, can provide evidence of students' perceived development in communicative competence, critical thinking, and classroom participation. Qualitative data, such as teacher observations, student reflections, and interview responses, can offer deeper insights into how debate-based pedagogy influences students' use of English, reasoning processes, and engagement. Mixed-methods research is useful when a study seeks to combine numerical trends with qualitative explanations of participants' experiences (Creswell & Plano Clark, 2018).

The study is therefore designed as a small-scale classroom intervention to examine how debate-based pedagogy can support Vietnam's transition from EFL-oriented instruction to an ESL-oriented learning environment. Rather than treating debate as a competitive extracurricular activity, the intervention positions debate as a communicative classroom pedagogy through which students use English to express opinions, construct arguments, evaluate evidence, respond to opposing views, and participate in meaningful interaction.

3.2 Participants

The participants are Vietnamese English-language learners aged approximately 8 to 15, studying in English-language classes where debate-based activities are integrated into instruction. The learners may come from primary and lower-secondary levels, with varying levels of English proficiency. Because the age range is broad, the intervention and research instruments are adapted to students' developmental stage and language ability.

The research also includes data from teachers, including debate teachers and English language teachers. Teacher data are important because the study not only investigates students' learning outcomes but also examines the pedagogical feasibility of debate-based instruction in Vietnamese English classrooms. Teachers' perceptions can provide valuable insights into students' communicative development, critical thinking, classroom participation, and the support required for effective debate integration.

3.3 Data Collection

Multiple data collection tools are used to strengthen the study's credibility. The use of multiple data sources allows the researcher to examine the intervention from different perspectives, including students' self-perceptions, teacher observations, classroom performance, and reflective responses.

3.3.1 Student questionnaires

Student questionnaires are used to collect learners' perceptions of how debate-based activities influence their English communication, thinking skills, confidence, and classroom participation. Because the participants may include students aged 8-15, two versions of the

questionnaire are used. For students aged 8-11, the questionnaire uses simple statements and a three-point response scale. For students aged 12-15, the questionnaire uses a five-point Likert scale and includes more detailed items related to speaking, argumentation, interaction, confidence, and challenges. Although the Skittles group and Juniors group differ in age, language proficiency, and cognitive development, the questionnaires are designed around the same core constructs: communicative competence, critical thinking, English-rich classroom interaction, confidence, and perceived challenges. The wording, response scale, and task complexity are adapted to each age group, but the underlying constructs remain consistent. For younger learners, items are simplified and supported by clearer language and a three-point scale. For older learners, the five-point Likert scale allows more detailed responses. This design allows age-appropriate data collection while maintaining cross-group conceptual alignment.

Using age-appropriate questionnaire design is important in second-language research because questionnaire items need to be understandable and appropriate to respondents' language proficiency and cognitive level (Dörnyei & Taguchi, 2010). The questionnaire focuses on three main constructs: communicative competence, critical thinking, and English-rich classroom interaction.

3.3.2 Teacher questionnaires

Teacher questionnaires are used to collect data from debate teachers and English language teachers. The teacher questionnaire examines teachers' familiarity with debate, prior use of debate in English classes, understanding of debate as pedagogy, perceptions of communicative competence, views on English as a tool for learning and thinking, and perceived implementation challenges. Teacher data are important because teachers can provide professional observations of students' language use, classroom participation, and the development of critical thinking. The teacher questionnaire also helps identify training needs and practical implications for English language teaching in Vietnam.

3.3.3 Classroom observation notes

Classroom observation notes are used to document students' participation and language use during debate-based activities. The researcher or teacher-observer records evidence of student talk time, peer-to-peer interaction, use of English during group work, willingness to speak, use of reasons and examples, listening and response, respectful disagreement, spontaneous English use, and participation of quieter students. Observation data are particularly useful for examining whether debate creates a more English-rich classroom environment.

3.3.4 Teacher interviews or focus groups

Short teacher interviews or focus groups may be conducted to gain deeper qualitative insights. Interview questions can explore how teachers understand the EFL-to-ESL transition, how debate differs from ordinary speaking activities, whether debate supports communicative competence, how debate supports critical thinking, what challenges students face, what scaffolding strategies are useful, and what training teachers need. Teacher interviews help explain and enrich the questionnaire and observation data.

3.4 Ethical Considerations

Because the study involves young learners, ethical considerations are especially important. Participation should be voluntary, and students should be informed that their responses will be used only for research purposes. Where necessary, parental or institutional

consent should be obtained. Students' names should not be used in the paper, and all data should be anonymised.

Students should also be reassured that the questionnaire is not a test and that there are no right or wrong answers. This is important because young learners may otherwise feel pressure to give positive responses. The teacher-researcher should make clear that honest answers will help improve future English lessons.

4. Data Analysis

4.1 Quantitative Analysis

4.1.1 Student survey analysis

For the Skittles group, the strongest results appeared in communication and early thinking skills. The overall positive response rate was 66.3% for communication-related items and 63.7% for thinking-related items. The strongest individual item was "When I gave an idea, I tried to explain my reason in English," with 90% of students selecting "Yes." Another strong item was "I tried to give an example to make my idea clearer," with 82.5% selecting "Yes." These results suggest that even younger learners were beginning to use English not only to speak, but also to explain, justify, and support their ideas.

However, the Skittles group showed weaker results in affective and motivational areas. Only 40.0% of students selected "Yes" for feeling brave when speaking English during debate activities. Similarly, 42.5% said debate made English class more enjoyable, and 46.3% said they would like to do debate activities again in the future. This suggests that debate can support young learners' communication and reasoning, but it may also feel emotionally demanding if not sufficiently scaffolded.

For the Juniors group, the strongest construct was critical thinking, with a mean score of 3.94 out of 5 and 69.6% of responses were positive. Communicative competence was also strong, with a mean of 3.87 and 67.1% of responses being positive. English-rich classroom interaction had a moderate mean of 3.68, while confidence, motivation, and learner agency had a lower mean of 3.48. The challenge section had a mean of 3.68, showing that students also recognised significant difficulties in debate-based learning.

When it comes to critical thinking, one thing really stands out - debate is a powerful tool. The top-rated activity for younger students was a statement that said, "Debate helped me understand that strong opinions need reasons. "A lot of students agreed with this, with an average rating of 4.14 and over 81% of responses being positive. What's more, students also liked using examples to support their arguments, developing counterarguments, responding to opposing views, and examining topics from different angles. This shows that using debates in the classroom can really help students move away from just expressing their opinions and towards communicating their thoughts in a more thoughtful and reasoned way. By doing this, students can develop stronger critical thinking skills and learn to express themselves more effectively.

4.1.2 Teacher survey analysis

When we looked at what teachers thought about using debates in the classroom, we found that they really liked the idea. They thought it was a great way to help students develop their critical thinking skills - almost 96% of teachers agreed with this, and they rated it an average of 4.63 out of 5. They also thought that English was an important tool for learning,

thinking, and communicating, with over 90% of teachers agreeing and an average rating of 4.57.

Teachers were also very supportive of using debates to teach, with an average rating of 4.54 and over 92% agreeing it was a good idea. Additionally, they thought that debates could help students become better communicators, with an average rating of 4.51 and almost 86.7% of teachers agreeing. Overall, teachers seemed to think that debates were a great way to help students learn and develop important skills.

Teachers strongly agreed that English should be used as a tool for learning, thinking, and interaction, not only as a school subject. This was the highest-rated item in the teacher questionnaire, with a mean of 4.83. They also strongly agreed that debate helps students express opinions, give reasons, and respond to others, and that debate helps students practise both language skills and thinking skills. Both items had a mean of 4.78.

At the same time, the teacher data revealed a gap between belief and practice. Although teachers viewed debate very positively, debate was not yet a regular classroom practice for many respondents. Only 17% reported using debate often or very often. By contrast, 39.1% used debate only sometimes, 30.2% had used it once or twice, and 13.2% had never used debate before. This shows that even though teachers like the idea of debate, they need more help, resources, and confidence to make it a regular part of their teaching. They need to feel more comfortable using debate in class and have the tools to do it well.

4.2 Qualitative Analysis

When we look at open-ended responses from both students and teachers, some common themes emerge. Many students said they felt more confident, learned new words, practised speaking more, and improved at organising their thoughts. They also said they improved at listening to others and giving reasons for their views. But they also had some difficulties, including limited vocabulary, finding ideas, responding quickly, giving evidence, and building rebuttals. Some students got nervous, had trouble with pronunciation, and felt like they were running out of time. To make things better, students wanted more help with vocabulary, some examples to get them started, and feedback from their teachers. They also wanted more time to prepare and some easier or more interesting topics to debate.

Teacher open-ended responses showed similar patterns. They viewed debate as useful for developing communication, critical thinking, argument structure, confidence, and collaboration. They also identified challenges such as limited vocabulary, low student confidence, large class sizes, lack of time, an exam-oriented classroom culture, and insufficient teacher training. Many teachers requested practical workshops, sample lesson plans, model banks, sentence frames, simplified rubrics, classroom management strategies, and opportunities to observe effective debate lessons.

Therefore, the qualitative data helped us better understand the quantitative findings. Both students and teachers thought debate was a good thing, but both groups emphasised that it requires scaffolding. This is an important finding because it suggests that debate-based pedagogy should not be treated as an unstructured speaking activity. It needs to be designed as a guided learning process.

While a fuller thematic analysis may be developed in future research, the present summary provides sufficient qualitative evidence to interpret the main quantitative findings and clarify the practical conditions needed for classroom implementation.

4.2.1 Student perspectives: confidence, language limits, and argument development

Student open-ended responses showed that debate helped learners practise speaking, organise ideas, listen to classmates, and explain opinions with reasons. Many students reported that debate encouraged them to use English more actively because they had to express a position, give examples, and respond to others. However, students also identified several barriers. The most common challenges were limited vocabulary, difficulty finding ideas quickly, nervousness when speaking, pronunciation problems, limited preparation time, and difficulty constructing rebuttals. These responses help explain why the quantitative results were stronger for communication and reasoning than for confidence and motivation.

4.2.2 Teacher perspectives: debate as useful but difficult to sustain

Teacher responses confirmed that debate was viewed positively as a tool for communication, critical thinking, collaboration, confidence-building, and argument structure. However, teachers also emphasised that debate is difficult to implement regularly without practical support. They identified large class sizes, limited lesson time, rigid curricula, exam-oriented assessment, uneven student proficiency, and insufficient teacher training as major barriers. This explains the belief-practice gap found in the quantitative data: teachers valued debate highly, but many did not yet feel prepared to use it frequently.

4.2.3 Scaffolding as the central condition for success

Across both student and teacher responses, scaffolding emerged as the central condition for successful debate-based pedagogy. Students requested vocabulary support, model answers, clearer examples, more preparation time, and easier or more engaging topics. Teachers similarly requested sentence frames, simplified rubrics, classroom management strategies, model lesson plans, debate motion banks, and opportunities to observe effective debate lessons. These findings suggest that debate should not be introduced as an open-ended speaking activity. Instead, it should be designed as a guided learning cycle that includes preparation, language support, structured interaction, feedback, and reflection

5. Findings

5.1 Debate supported communicative competence beyond grammar accuracy

The first major finding indicates that debate-based pedagogy significantly enhances students' communicative competence by shifting the focus from grammatical accuracy to functional, reasoned language use. Both student and teacher data confirm that debate helps learners move beyond short, isolated answers toward more extended, meaningful communication.

- Skill Development: Students reported marked improvements in key areas, including clearly expressing opinions, organising ideas, providing supporting examples, and responding to classmates. Notably, 82.5% of students actively used examples to clarify ideas, and 70% reported producing longer, more complex English sentences.

- Teacher Validation: Teachers strongly agreed that debate enables students to develop essential competencies beyond grammar, including discourse organisation, interactional competence, and the ability to use functional language to agree, disagree, clarify, and politely challenge viewpoints.

- A Pathway to Spontaneous Communication: While the results are encouraging, the data also highlight that natural fluency is a gradual outcome. Only approximately 50–54% of junior students felt debate helped them achieve natural fluency or move entirely beyond

exercise-based learning. This suggests that debate functions best as a structured pathway toward spontaneous communication rather than an immediate guarantee of naturalistic fluency.

This finding supports the study's central argument that debate-based pedagogy serves as a practical, classroom-level intervention for Vietnam's EFL-to-ESL transition by fostering the structured, reasoned interaction necessary for real-world English use.

5.2 Critical Thinking was the strongest shared outcome across students and teachers

The second major finding is that critical thinking was the strongest and most consistent outcome across both student and teacher data.

- Consistent Success: For junior students (ages 12–15), critical thinking yielded the highest construct mean (3.94), with students reporting greater ability to provide evidence, build logical rebuttals, and evaluate multiple perspectives rather than simply expressing agreement or disagreement.

- Developmental Appropriateness: Even among younger "Skittles" learners (ages 8–11), the pedagogy successfully introduced foundational argumentation habits, such as providing reasons, using examples, and responding respectfully to opposing views.

- Teacher Consensus: Educators overwhelmingly identified critical thinking as a primary benefit, with a 4.63 mean rating and 94.3% agreement. They observed that debate helps students move from rote memorisation to higher-order cognitive tasks like argument construction, evidence evaluation, and perspective-taking.

This triangulation of data demonstrates that debate-based pedagogy is uniquely powerful because it integrates language production with cognitive processing. By requiring students to use English to reason, justify, and evaluate, debate transforms the language from a static subject of study into a dynamic tool for communication and inquiry. This directly supports the study's central thesis: debate facilitates the transition toward ESL-oriented environments by making English functional, reasoned, and cognitively demanding.

5.3 Debate helped move English from a subject to a tool for meaningful use

The third major finding is that debate supported the movement from learning English as a subject toward using English as a tool for communication, thinking, and interaction.

This finding is especially relevant to the study's focus on Vietnam's transition from EFL to ESL.

- Student Perception: A majority of Junior students (approx. 72%) reported that debate helped them learn English through actual use, thinking, and communication. Furthermore, 68.4% indicated that the approach enabled them to discuss real-life topics, moving beyond traditional textbook exercises.

- Teacher Alignment: Educators strongly aligned with this goal, assigning a mean rating of 4.83 to the principle that English should serve as a tool for interaction rather than just a subject. They acknowledged that debate provides a practical, classroom-level pathway to achieve this, aligning with the "English-as-a-medium" principle of CLIL and EMI by using the language for argumentation, inquiry, and reasoning.

- Limitations of Interaction: Despite these benefits, the data indicate that debate only partially creates an English-rich environment. While students succeeded in structured tasks, fewer than 42% felt comfortable initiating spontaneous peer interaction or asking questions in English.

In conclusion, this finding establishes that while debate is a powerful tool for fostering reasoned communication, it should be recognised as a structured strategy rather than an immediate substitute for full implementation of CLIL or EMI. To enhance its impact, students require more explicit support to develop the confidence and skills necessary for spontaneous, naturalistic classroom interaction.

5.4 Debate was seen as a classroom pedagogy, not merely a competition

The fourth finding reveals that teachers predominantly conceptualise debate as a versatile classroom pedagogy rather than merely a competitive extracurricular activity. Survey data indicates that the vast majority of teachers utilise debate for speaking (79.2%), critical thinking (66.0%), discussion strategies (52.8%), and authentic communication (43.0%), with only 13.2% viewing it primarily as a competition. Educators strongly support adapting debate into short, flexible classroom tasks—such as pair or group arguments—to foster reasoned interaction, dismissing the notion that it must always take a formal, competitive form. This finding suggests that teachers are more likely to implement debate when it is presented as a flexible classroom routine rather than as a formal competition. In practical terms, this includes short formats such as pair debates, opinion-line discussions, quick rebuttal exchanges, role-card debates, and three-minute group debates. These formats preserve the essential features of debate while reducing pressure on lesson time, classroom management, and student anxiety. Despite this positive perception, a significant belief-practice gap exists: while teachers value the approach, only 17.0% implement it frequently. This disparity appears to be rooted in moderate teacher confidence, lack of training, resource limitations, and uncertainties regarding assessment, indicating that while debate is recognised as a valuable pedagogical tool, it has yet to be fully embedded into routine teaching practice.

5.5 Scaffolding is the main condition for successful debate-based pedagogy

The fifth finding emphasises that scaffolding is the critical prerequisite for successful debate-based pedagogy. The data consistently show that while debate is effective, its success is entirely dependent on providing students with robust linguistic and emotional support.

- Student Challenges: Junior students identified limited vocabulary as their primary barrier (nearly 76%), followed closely by performance anxiety, difficulty generating ideas, rebuttal challenges, and the pressure of spontaneous response. Similarly, younger "Skittles" learners showed low emotional confidence, with only 40% feeling "brave" while debating, highlighting a need for more playful and guided formats.

- Teacher Consensus: Teachers unanimously agreed that systematic pedagogical design—not just enthusiasm—is necessary. They require clear structures, role assignments, time control, and language supports to mitigate classroom barriers such as large class sizes and time constraints.

- Systematic Support: Findings indicate that to move beyond "unstructured speaking" into reasoned communication, teachers must prioritise pre-debate scaffolding, including vocabulary lists, sentence frames, argument templates, and clear assessment rubrics.

Examples of scaffolding tools include sentence frames such as “I agree with this because…,” “One reason is…,” “For example…,” “I understand your point, but…,” “My evidence is…,” and “Therefore, I believe…”. Argument templates can also support students by guiding them through a clear sequence: Claim → Reason → Example/Evidence → Explanation → Link back to the motion. Similarly, rebuttal templates can help students organise responses through the structure of Opposing idea → Weakness/problem → Counter-

reason → Example → Conclusion. A simplified rubric may assess students on clarity of opinion, use of reasons, use of examples, response to others, respectful interaction, and English effort

This finding has direct implications for teacher training in Vietnam. To effectively support the transition from EFL to ESL, professional development must move beyond teaching "debate formats." Instead, training must equip educators with the practical skills to scaffold language, manage participation, foster emotionally safe environments, and adapt complex debate structures for diverse age groups and proficiency levels. Without this intentional design, the benefits of debate remain largely inaccessible

6. Recommendations

Based on the findings, debate-based pedagogy should be implemented as a structured, scaffolded, and age-appropriate approach rather than as an unplanned speaking activity or a formal competition only. The recommendations below are directed at teachers and educators, educational managers, and policy makers.

6.1 Recommendations for Teachers and Educators

- Prioritise Scaffolding: Implement intentional language and structural support (e.g., sentence frames, argument templates, and vocabulary sheets) to reduce cognitive load and help students organise ideas.
- Create an Emotionally Safe Environment: Mitigate speaking anxiety by starting with low-stakes, manageable tasks such as pair work, small-group rehearsals, and mini-debates, and focus feedback on effort and reasoning rather than just grammatical accuracy.
- Differentiate by Level: Tailor activities to age and proficiency; use playful, simplified tasks for younger learners and gradually introduce complex argumentation and rebuttal techniques for older students.
- Teach Spontaneous Interaction: Explicitly practice question stems and clarification strategies to help students move from prepared speech toward natural, spontaneous classroom interaction.
- Adopt Learning Cycles: Utilise post-debate reflections, where students identify successes and areas for improvement, and teachers provide delayed feedback to reinforce language and thinking skills.

6.2 Recommendations for Educational Managers

- Develop Practical Resources: Close the belief-practice gap by providing ready-to-use materials such as motion banks, lesson plans, and simplified assessment rubrics.
- Promote Flexible Formats: Normalise debate as a routine classroom activity—such as 10–15 minute tasks—rather than reserving it only for special events or formal competitions.
- Navigate Constraints: Integrate debate into existing curricula by embedding it into reading, writing, and speaking outcomes to ensure it fits within time and resource constraints.
- Foster Peer Learning: Organise model lessons and internal workshops to demonstrate how debate can work in real, large-class settings.

6.3 Recommendations for Policy Makers

- Formally Recognise Debate: Integrate debate-based pedagogy into national guidelines as a practical pathway to support Vietnam's transition from EFL to ESL, with a focus on English as a functional tool for reasoning and communication.
- Targeted Professional Development: Focus teacher training on scaffolding language, managing classroom participation, assessing higher-order thinking, and adapting debate for diverse proficiency levels.
- Broaden Assessment Frameworks: Reform English assessment to incentivise debate by including criteria such as discourse organisation, interactional competence, use of evidence, and reasoned rebuttal, rather than focusing solely on grammatical accuracy and rote exam performance.

Overall, the successful transition to ESL-oriented English education in Vietnam depends on ensuring that debate is viewed not as an "unplanned speaking activity," but as a structured classroom approach that transforms English into a tool for cognitive and communicative development.

7. Research limitations

This study has several limitations. Firstly, due to space constraints, this paper presents a summary of the qualitative findings. A comprehensive thematic analysis of interview transcripts and observation notes is reserved for future publication. However, the qualitative evidence presented in this paper is sufficient to explain the main quantitative patterns and identify the practical conditions required for successful debate-based pedagogy.

Secondly, external pressures such as large class sizes, rigid curricula, and exam-oriented assessment were identified as significant barriers but were not experimentally isolated. Therefore, this study cannot determine the separate effect of each constraint on the success of debate-based pedagogy. Instead, the findings should be interpreted as classroom-based evidence from a real Vietnamese EFL context where these constraints interact with teacher practice.

Lastly, the study involved 159 students and 53 teachers, making it a small-scale action research project rather than a nationally representative study. For this reason, the findings should be understood as context-specific insights rather than broad generalisations about all Vietnamese schools. However, the findings may be transferable to other Vietnamese EFL contexts that share similar features, such as grammar-oriented instruction, limited student speaking opportunities, exam pressure, mixed proficiency levels, and growing policy interest in English-rich learning environments.

Future research could examine the long-term effects of debate-based pedagogy, compare debate implementation across different regions or school types, and investigate how specific constraints such as class size, curriculum flexibility, and assessment culture influence classroom debate outcomes.

8. Conclusion

This study suggests that debate-based pedagogy can offer a practical, classroom-level pathway for supporting Vietnam's ESL-oriented policy goals, particularly in EFL classrooms where students need more opportunities to use English for communication, reasoning, and interaction. By engaging students in argumentation, inquiry, and evidence-based reasoning, educators can create "English-rich" environments even in contexts where full CLIL or EMI implementation remains challenging. Debate does not require a complete overhaul of the

curriculum; rather, it functions as a flexible strategy that, when integrated into routine practice, fosters the kind of reasoned, spontaneous interaction essential for Vietnam's international integration.

The transition toward an ESL-oriented education system in Vietnam is a pedagogical journey that requires more than policy directives; it requires a shift in classroom culture. Debate-based pedagogy, when supported by systematic teacher training, practical resources, and a broadening of assessment frameworks beyond grammar, provides a tangible method for achieving this. Because this was a small-scale action research project, the findings should not be treated as nationally representative. However, they offer context-specific insights that may be adapted by teachers and institutions facing similar classroom realities, including limited speaking opportunities, exam pressure, mixed proficiency levels, and the need for practical English-rich activities. Future research could explore the longitudinal impact of these practices and examine how debate can be better adapted to support spontaneous peer interaction, helping learners move beyond the task's structure toward the natural fluency required for real-world English communication.

References

Aclan, E. M., & Aziz, N. H. A. (2015). Exploring parliamentary debate as a pedagogical tool to develop English communication skills in EFL/ESL classrooms. *International Journal of Applied Linguistics and English Literature, 4*(2), 1–16. https://doi.org/10.7575/aiac.ijalel.v.4n.2p.1

Aclan, E. M., Aziz, N. H. A., & Valdez, N. P. (2016). Debate as a pedagogical tool to develop soft skills in EFL/ESL classroom: A qualitative case study. *Pertanika Journal of Social Sciences & Humanities, 24*(1), 213–240.

Akerman, R., & Neale, I. (2011). *Debating the evidence: An international review of current situation and perceptions*. Akerman, R., & Neale, I. (2011). Debating the Evidence An International Review of Current Situation and Perceptions. The English-Speaking Union. - References - Scientific Research Publishing

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–27). Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47. https://doi.org/10.1093/applin/1.1.1

Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics, 6*(2), 5–35. https://doi.org/10.5070/L462005216

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Dearden, J. (2014). *English as a medium of instruction: A growing global phenomenon*. British Council.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). Routledge.

El Majidi, A., de Graaff, R., & Janssen, D. (2024). Debate as a pedagogical tool for developing speaking skills in second language education. *Language Teaching Research, 28*(6), 2431–2452. https://doi.org/10.1177/13621688211050619

Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.

Iman, J. N. (2017). Debate instruction in EFL classroom: Impacts on the critical thinking and speaking skill. *International Journal of Instruction, 10*(4), 87–108. https://doi.org/10.12973/iji.2017.1046a

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer Singapore. https://doi.org/10.1007/978-981-4560-67-2

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245–1260. https://doi.org/10.1111/1467-8624.00605

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Thi Thu Hien, T. (2022). Teacher's changes and challenges in CLIL implementation. *VNU Journal of Foreign Studies, 38*(6), 44–58. https://doi.org/10.25073/2525-2445/vnufs.4883

Dörnyei, Z., & Dewaele, J.-M. (2022). Questionnaires in Second Language Research: Construction, Administration and Processing (3rd ed.). New York: Routledge – Taylor & Francis. ISBN: 9781032364339.

https://www.researchgate.net/publication/363746272_Questionnaires_in_Second_Language_Research_Construction_Administration_and_Processing

Ministry of Education and Training of Viet Nam. (2025, October 29). Ban hành Đề án "Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học giai đoạn 2025-2035, tầm nhìn đến năm 2045" [Promulgating the project "Making English the second language in schools for 2025–2035, with a vision to 2045"]. URL: https://moet.gov.vn/tin-tuc-cot-phai/chuong-trinh-giao-duc-pho-thong/ban-hanh-de-an-dua-tieng-anh-thanh-ngon-ngu-thu-hai-trong-truong-hoc-giai-doan-2025-2035-tam-nhin-den-nam-2045-.html

Prime Minister of the Government of Viet Nam. (2025, October 27). Quyết định số 2371/QĐ-TTg: Phê duyệt Đề án "Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học giai đoạn 2025 - 2035, tầm nhìn đến năm 2045" [Decision No. 2371/QĐ-TTg approving the project "Making English the second language in schools for 2025–2035, with a vision to 2045"]. Government Portal. URL: https://vanban.chinhphu.vn/?docid=215728&pageid=27160 PDF: https://datafiles.chinhphu.vn/cpp/files/vbpq/2025/10/2371-ttg.signed.pdf



HUFLIT INTESOL
INTERNATIONAL
CONFERENCE 2026