**Exploring the Effects of AI-Powered IELTS Writing Checker Bot on Grammar Competency among IELTS Learners**

*Nguyen Thi Huyen Tran*

**Abstract**

In the technological era, integrating AI tools into teaching writing skill for EFL learners has proved effective and received considerable attention of many educators. This study aims at exploring the impact of **IELTS Writing Checker Bot**, an AI-powered chatbot designed to evaluate IELTS essays, on improving **grammar competency** among IELTS learners. The research employed a **mixed-method experimental design**, involving 80 senior students at Ho Chi Minh City University of Foreign Languages and Information Technology. The experimental group utilized the IELTS Writing Checker Bot within a period of 12 weeks while the control group received traditional instructions. Quantitative data were collected through pre-tests, post-tests, and questionnaires. The results reveal that the experimental group outperformed their peers in the control group in terms of the "Grammatical Range and Accuracy" criterion. In addition, according to the results of the questionnaire and interviews, students characterized the IELTS Writing Checker Bot as a user-friendly **and inspiring tool** which made it easier for them to correct errors and encouraged them to practice writing essays more frequently. These findings suggest that IELTS Writing Checker Bot can be an effective tool to enhance grammar learning as well as improving writing skill. Further research is recommended to investigate its effectiveness when applied to a larger group of EFL students.

***Key words***: IELTS Writing Checker Bot, AI-powered Chatbot, grammar competency, IELTS writing, IELTS learners, EFL learners.

1. **Introduction**

Writing plays a crucial role in language education, promoting critical thinking, self-expression, and effective communication (Hyland, 2003). For English language leaners, writing plays a significant role in develop capabilities in various areas including organization, coherence, grammar and vocabulary. Therefore, effective writing assessment in English language programs is essential for measuring students' language proficiency and ensuring they meet program learning outcomes. In this regard, formative feedback serves as a continuous process which provides timely support for students, including giving feedback and guidance on content, organization, grammar, vocabulary and style (Zhu et al., 2020). However, implementing this task in a way that meets every student’s individual feedback needs poses a considerable chalenge as it is both time-consuming and impractical, especially in large-size classes (Golzar et al., 2022).

As traditional feedback methods become less effective, researchers emphasize the need for smarter and more personalized approaches (Kohnke et al., 2023). The rise of AI-powered tools like Grammarly, QuillBot, Wordtune, and ChatGPT has made it more accessible for students to receive writing feedback. These technologies have significantly improved automated writing evaluation and feedback processes (Zhao, 2022; Marzuki et al., 2023).

As a powerful generative language model, ChatGPT holds significant potential to support student writing. Compared to traditional automated writing evaluation (AWE) tools, it offers more personalized and student-friendly assistance (Guo et al., 2022; Rudolph et al., 2023). The integration of ChatGPT can fulfil instant feedback needs of students, aiding ideas generation, organization, ensuring accuracy and vocabulary appropriateness (Tai et al., 2023). Additionally, ChatGPT helps students overcome writer’s block, and boost confidence and motivation (Song et al., 2023). However, despite the potentials and effectiveness of ChatGPT for EFL students, these have been some concerns raised among scholars about students’ over-reliance on AI, which possibly leads to plagiarism risks, a lack of critical thinking ability and problem-solving (Song et al, 2023; Yan, D., 2023).

However, when the researcher allowed students to use ChatGPT as an AI-assisted writing tool in International English Language Testing System (IELTS) academic writing courses, it was challenging to determine whether the submitted take-home essays were entirely the students' own work or influenced by AI-generated content, making it hard to assess the student’s actual writing ability and progress. In addition, the lecturer was unable to observe the students’ practice process with ChatGPT or evaluate how well they applied the feedback to revise and improve their writing. Beyond the issue of tracking the extent to which the submitted essays reflected students’ own work and the level of revision with AI tools, another challenge identified by the researcher was helping students improve their grammatical accuracy within the limited time frame of the course as IELTS writing classes often prioritized essay structure and format and grammar instruction was not the primary focus. Although students were aware of common mistakes they frequently made in their essays, they had difficulty avoiding those mistakes due to the limited feedback opportunities provided by the lecturer throughout the course.

Therefore, to address these issues, the researcher developed IELTS Writing Checker Bot (IWCB), a customized Chatbot, to integrate it into the writing lessons as a tool to help students revise their IELTS essays. It provides automated feedback across four scoring criteria including Task Response, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy, help students identify errors, receive improvement suggestions, and revise their writing independently. In addition, the researcher shifted the focus from assessing the perfection of the submitted essays to evaluating the progress students made, so students were evaluated based on their revision process with IWCB and how well they responded to the feedback given by IWCB. This aims to encourage students to engage more actively in improving their writing, maintaining grammar accuracy and reduce the risk of plagiarism.

While numerous studies have hightlighted and confirmed the effectiveness of integrating AI-assisted tools and ChatGPT in the process of teaching academic writing skill, there has been limited investigation into the integration of a specialized GPT Chat Bot programmed to give scores and feedback for IELTS essays and its impact on grammar competency among IELTS learners.

This study will address two research questions:

1. To what extent does the employment of IWCB improve grammar competency among IELTS learners?

2. What are students’ perceptions towards the impact of IWCB on their grammar competency?

By examining the effectiveness of IWCB in grammar competency and students’ perceptions, this study will contribute to the area of AI-assisted tools in language learning and provide valuable insights into the use of AI-powered Writing Checker Bot in IELTS writing instructions.

**2. Literature Review**

***2.1 Chatbots and AI Role in English Language Learning***

Chatbots are computer programs that can make conversations with humans (Haristiani, 2019). They can be incorporated with artificial intelligence (AI) to provide information, answer questions and perform various tasks (Skrebeca et al., 2021). Thanks to AI technologies like machine learning and natural language processing (NLP), chatbots can understand and respond to human language (Johri et al., 2021). In language learning, Chatbots using AI are powerful tools because they offer interactive and immediate feedback as well as peronalized practices (Kamalov et al., 2023). Learners can receive instant individualized feedback, which helps them identify and correct their mistakes in real-time, and thus focus on the areas they need to improve (Almusaed et al., 2023).

Given the advancement in technology, the 24/7 availability of Chatbots enable learners to practice at their own pace and take control of their learning journey (Hew et al., 2023). Furthermore, Chatbots can be customized to suit individual learning styles, making the learning experience more tailored and personalized (Kooli, 2023). This can encourage students to keep practicing towards achieving their goals and become more motivated.

***2.2 AI-powered Chatbots to support writing feedback***

Many studies on AI-powered chatbots have revealed their advantages and disadvantages as writing assistants within EFL writing instruction.A review study of Barrot’s (2023) provided valuable insights on ChatGPT potentials, emphasized that this tool possesses numerous benefits such as offering immediate, personalized feedback tailored to students’ writing progress. It also highlighted the promising potentials of ChatGPT as a virtual tutor and language input provider. However, some academics raised concerns regarding academic integrity while using AI.Similarly, Song et al. (2023) employed the mixed-method approach to examine the impact of AI-assisted writing instruction using ChatGPT on academic writing skills among 50 Chinese EFL university students learning in an IELTS writing course as well as exploring their motivation. The quantitative results underscored the marked enhancement in academic writing skills and motivation among the experimental group, particularly in organization, coherence, grammar, and vocabulary. Moreover, qualitative data further showed heightened motivation, improved confidence and self-efficacy. However, some concerns regarding contextual suitability of AI-generated feedback and the risk of students becoming overly dependent on AI were discussed. Another research conducted by Yan, D. (2023) shared the same positive influence of ChatGPT on student’s writing skills and the method of teaching academic writing. It revealed the effect of ChatGPT’s text generation function on students’ writing skills in a one-week L2 writing practicum and used qualitative approach to investigate students’ reflections towards ChatGPT on writing performance. The results indicated that ChatGPT exerted potent effects on writing pedagogy. However, the issues of academic integrity and fairness were also voiced by students.

Some other researches focused on the more specific functions of ChatGPT as an essay revising tool or an automated feedback tool. Han et al. (2023) integrated ChatGPT into second language (L2) writing courses for 213 undergraduate and graduate EFL students by developing a digital learning platform called RECIPE (Revising an Essay with ChatGPT) where ChatGPT acts as a personalized writing assistant. This platform serves the purpose of guiding students in revising and improving their essays step by step. He discovered that this structured interaction not only supported students in recalling key points from class lessons but also enabled them to receive more class-relevant feedback from ChatGPT. By the end of the course, students’ perspective towards ChatGPT was positive. Another research investigated the impact of ChatGPT as a formative feedback tool on undergraduate ESL students’ writing skills was conducted by Mahapatra (2024), adopting a mixed methods quasi-experimental design. 35 first year science and engineering students in an Indian university participated in the experimental group compared to 37 in the control group. Students learned 3 writing types including process, comparison and cause-effect for six hours using ChatGPT as an automated feedback tool. The study highlighted the strong positive effect of ChatGPT on students' academic writing performance, showing their favorable views of its use. It also positioned ChatGPT as an effective dialogic feedback tool which is promising in large writing classrooms. However, the limitations involve exploring the impact of ChatGPT on more writing genres and micro aspects of writing skills.

Taken together, the existing literature on AI-powered chatbots has made significant contributions to our understanding of their positive role as automated feedback tools, essay revising tools and writing-assisted tools. Numerous advantages of ChatGPT were discovered such as enhancing writing performance, increasing learner motivation, fostering self-regulated learning, and boosting confidence. However, the integration of ChatGPT in writing lessons also raised concerns about over-reliance on idea generation, academic integrity, plagiarism and the ability of maintaining the same level of improvement when not using AI. Additionally, current research has primarily focused on writing skill, with limited attention given to the micro-level components of writing. Furthermore, most chatbot applications incorporated in language learning contexts have centered on ChatGPT in general, showing a notable gap in research on a chatbot specializing in writing assessment. Therefore, with the aim of making the most of AI-supported chatbot as a writing-assisted tool in teaching IELTS writing in a short course and minimizing the potential risks related to AI use, the reseacher developed the IELTS Writing Checker Bot focusing on revising, correcting essays and giving IELTS scores based on four official criteria. The study then measured students’ improvements in grammar skill and explore their perceptions of using this platform.

3. Methodology

***3.1 Pedagogic settings and Participants***

The study was conducted at Ho Chi Minh City University of Foreign Languages and Information Technology. The participants of the study are 80 third-year EFL students from 2 classes doing a course in Advanced English skills, with the textbook being Mindset for IELTS 3, the highest level one among the 4-level Mindset for IELTS series published by Cambridge University Press publisher. These classes are intact groups which were offered to the researcher by the Department of Foreign Language in the semester 3, Intake 2024-2025. One class is the experimental group and the other class is the control group, with 40 students in each class. Both groups attended a 3-hour writing session every week over a 12-week course. The experimental group used IWCB tool to receive scores and feedback based on 4 writing criteria when completing the writing homework. Rather than being graded solely on their initial submission, students were required to revise and resubmit their essays multiple times through the Bot until there were no errors detected by IWCB. Once revisions were complete, students shared the final link generated by IWCB platform with the researcher to observe and assess the entire revising process. Meanwhile, the control group acquired the traditional assessment including the researcher’s scores and feedback for their essays.

***3.2 Instrumentation***

In this study, the IELTS Writing Checker Bot was developed using the OpenAI GPT-based platform to serve as a customized tool for assessing and giving feedback on IELTS writing Task 1. The setup process involved designing a structured prompt that clearly defined the bot’s role, functions, and feedback approach. The IELTS Task 1 Band Descriptor was uploaded to ensure consistent and accurate evaluation. The full prompt used to configure the IELTS Writing Checker Bot is provided in the Appendix.

***3.3 Design of the study***

The study adopted a mixed-method quasi-experimental design, which integrated quantitative data from pre-tests and post-tests with qualitative insights from questionnaires and interviews, providing a comprehensive assessment of IWCB’s impacts on students’ grammatical proficiency.

***3.4 Data collection***

*3.4.1 Pretest and Posttest*

In the study, a pre-test and post-test were administered to evaluate the effectiveness of the intervention, utilizing IELTS academic writing task 1. The essay questions in the pre-test and post-test were drawn respectively from the Cambridge Practice Test 17 and 18, published by Cambridge University Press. Students’ essays were marked based on the official IELTS task 1 band descriptor (see Appendix 1) focusing on the four criteria: task achievement, coherence and cohesion, lexical resources, and grammatical range and accuracy. Then the scores of “grammatical range and accuracy” criterion were extracted for statistical analysis. The pretest was conducted after students completed their first writing lesson, rather than before instruction in order to avoid biased results, which occur when students are unfamiliar with the task and thus are unable to produce a complete response. The posttest was administered at the end of the course. Both tests were conducted in paper-based format to minimize the risk of academic dishonesty. Furthermore, to ensure the objectivity of the scores, the writing samples of pretest and posttest were evaluated by two independent raters, including the researcher and another lecturer, both of whom are experienced IELTS instructors who had received official training annually at IDP Education, a global education organization and an official co-owner of the IELTS test. 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 | 0.903 | < 0.01 |
| Rater 2 | 0.903 | 1 |  |

**Table 2**

*Correlation between post-test scores given by two raters (n = 80)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Rater 1** | **Rater 2** | **Sig. (2-tailed)** |
| Rater 1 | 1 | 0.953 | < 0.01 |
| Rater 2 | 0.953 | 1 |  |

The Pearson correlation coefficients for both the pre-test and post-test scores between the two raters are greater than 0.8 (p < 0.001). This indicates a very high level of correlation, demonstrating strong consistency and agreement between the two raters during the scoring process. Therefore, with the input data calculated by the average scores of the two raters, these results confirm the objectivity and reliability of the input data, ensuring that the subsequent analysis is not influenced by variations in rater judgments.

*3.4.2 Questionnaires*

To examine students' perceptions of the IELTS Writing Checker Bot, a questionnaire was designed on Google Form and consisted of three parts: Part 1 – Demographic information, Part 2 – Students’ Perceptions of the effectiveness of IWCB, Part 3 – Students’ motivation when using IWCB. The questionnaire was conducted online, and its reliability was measured using Cronbach’s alpha, as indicated in the following table.

**Table 3**

***Reliability of the Questionnaire (n = 80)***

|  |  |  |
| --- | --- | --- |
| **Part** | **Number of items** | **Cronbach’s Alpha** |
| Part II: Effectiveness of IWCB | 15 | 0.94 |
| Part III: Motivation | 5 | 0.87 |

The reliability test results showed that the Cronbach's alpha values were greater than 0.7. This indicates that the questionnaire has good reliability and high internal consistency, which is appropriate for use in the research.

*3.4.3 Interview*

The purpose of the interview was to explore learners' general attitude towards using IWCB for writing practice, with particular focus on its impact on grammar skills. A structured interview format was employed, in which the researcher followed a fixed set of questions in a specific order. Five students were randomly chosen to participate, coded as S1 through S5, and one-on-one interviews were implemented online via Microsoft Teams to maintain confidentiality and create a comfortable environment for students to share their ideas.

**4. Findings and discussions**

***4.1 Grammar Pretest and Posttest***

*4.1.1 Independent sample t-test of pretest results.*

A descriptive statistics and an Independent samples t-test were conducted to analyze the pre-test scores of the criterion “grammartical range and accuracy” between the the control group (CG) and experimental group (EG) as illustrated in the following table. Prior to the test, the normal distribution of the data was confirmed by examining the Normal Q-Q plots, which showed that the scores in both groups were evenly distributed, meeting the assumption of normality.

**Table 4**

*Descriptive Statistics of the Groups (n = 80)*

|  |  |  |
| --- | --- | --- |
| **Group** | **Sample size** | **Mean ± Standard Deviation** |
| CG | 40 | 4.78 ± 0,54 |
| EG | 40 | 4.64 ± 0,47 |

**Table 5**

*Independent t-test results for Pre-test scores between the two groups (n = 80)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Mean Difference** | **t** | **df** | **p** | **95% Cl** |
| **Equal variances assumed** | 0.14 | 1.22 | 78 | 0,23 | (-0.09; 0.36) |

*\* Levene's test for equality of variances: F = 1.16, p = 0.29 (> 0,05), variances are assumed equal).*

The result of Levene’s test showed a p-value of 0.29 (> 0.05), indicating that the variances between the two groups were homogeneous. Therefore, the condition for conducting a t-test with the assumption of equal variances was met.

The independent t-test analysis yielded a t-value of 1.22 with degrees of freedom (df) = 78 and a significance level of p = 0.23. Since the p-value is greater than 0.05, no statistically significant difference was found in the pre-test scores between the EG (mean = 4.64; standard deviation = 0.47) and the CG (mean = 4.78; standard deviation = 0.54). This suggests that the participants in both groups had equivalent proficiency levels in grammar skill at the beginning of the study.

*4.1.2 Independent sample t-test of posttest results.*

An independent samples t-test was conducted to compare the post-test scores of the criterion “grammartical range and accuracy” between the CG and the EG in order to evaluate the effectiveness of the new method. Prior to the analysis, the normal distribution of the data was confirmed through Normal Q-Q plots, which showed that the scores in both groups were normally distributed in accordance with the assumption of normality.

**Table 6**

*Descriptive Statistics of the Groups (n = 80)*

|  |  |  |
| --- | --- | --- |
| **Group** | **Sample size** | **Mean ± Standard Deviation** |
| CG | 40 | 4.99 ± 0.30 |
| EG | 40 | 5.51 ± 0.75 |

**Table 7**

*Independent t-test results for Pre-test Scores between the two groups (n = 80)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Mean Difference** | **t** | **df** | **p** | **95% Cl** |
| **Equal variances not assumed** | -0.52 | -4.1 | 52 | <0.01 | (-0.78, -0.27) |

*\* Levene's test for equality of variances: F = 24,1, p < 0,01, variances are not assumed equal).*

The Levene’s test yielded an F value of 24.1 and a p-value < 0.01, indicating that the variances between the two groups were not equal. Therefore, the t-test was conducted assuming unequal variances to ensure accurate results.

The test results revealed a statistically significant difference between the two groups in terms of post-test scores (t = -4.1; df = 52; p < 0.01). Specifically, the intervention group had a significantly higher average score than the traditional group (Mean Difference = -0.52; 95% CI: -0.78 to -0.27). This indicates that the EG achieved a more substantial improvement in learning outcomes compared to the CG. In other words, the analysis confirmed that the implementation of IWCB platform had a positive impact and significantly enhanced learner’s grammar competency compared to the traditional approach.

***4.2 Questionnaires***

**A total of 40 students from the experimental group participated in the online survey to assess their perceptions of using IWCB, including** 72% female and 28% male students. The mean scores and standard deviations (SD) for each item in the questionnaire are presented in the following table.

**Table 8**

*Students’ perceptions of using IWCB for grammar practice (n = 40)*

| **EFFECTIVENESS OF IWCB** | **MEAN** | **SD** |
| --- | --- | --- |
| After using IWCB, I have better control over my grammar accuracy | 4.35 | 0.67 |
| I can use tenses correctly and consistently | 4.35 | 0.67 |
| I can apply subject-verb agreement correctly | 4.22 | 0.77 |
| I can use comparatives accurately | 4.25 | 0.67 |
| I can use articles (a/an/the) correctly | 4.43 | 0.71 |
| I can use passive voice accurately | 4.28 | 0.68 |
| I can punctuate properly | 4.43 | 0.81 |
| After using IWCB, I can improve the grammar range in the essay | 4.28 | 0.82 |
| I can use relative clauses more frequently and correctly | 4.08 | 0.80 |
| I can construct compound sentences more frequently and accurately | 4.05 | 0.82 |
| I can use conditional sentences more frequently and correctly | 3.85 | 0.88 |
| Overall, I am more confident in my grammar skill when I write independently without the support of IWCB | 4.35 | 0.80 |
| IWCB helps improve my Task Achievement score | 4.40 | 0.71 |
| IWCB helps improve my Coherence and Cohesion score | 4.35 | 0.78 |
| IWCB helps improve Lexical resources score | 4.20 | 0.82 |
| **MOTIVATION** | **MEAN** | **SD** |
| I think IWCB is easy to use for writing practice | 4.80 | 0.57 |
| IWCB motivates me to practice IELTS writing | 4.13 | 0.76 |
| I’m not worried about becoming overly dependent on IWCB | 4.15 | 0.80 |
| I am satisfied with IWCB as a writing assistant | 4.45 | 0.67 |
| I am willing to continue using IWCB for writing learning | 4.60 | 067 |
| **AVERAGE** | 4.28 | 0.54 |

### Effectiveness of IWCB for Grammar Practice

Students expressed highly positive perceptions regarding the effectiveness and satisfaction of using IWCB for English grammar practice.

They particularly valued IWCB's impact on improving grammatical accuracy. Items related to the use of tenses, subject-verb agreement, comparison, articles, and passive voice scored between **4.22 and 4.43** on a 5-point Likert scale, with relatively low standard deviations (SD ranging from **0.67 to 0.82**). Notably, **the use of articles received the highest average score (Mean = 4.43)**, indicating a strong improvement in this aspect.

The use of complex grammatical structures such as relative clauses, compound sentences, and conditional sentences also received positive ratings, although the average scores for these items were slightly lower (**Mean ranging from 3.85 to 4.08**), suggesting these may require further instructional support.

Students also reported improvements in **Task Achievement**, **Coherence and Cohesion, and Lexical Resources** in their IELTS writing when using IWCB (**Means ranging from 4.20 to 4.40**), highlighting the comprehensive support the tool provides for academic writing skills.

### Motivation and satisfaction

Most students found IWCB easy to use for writing practice (**Mean = 4.68, SD = 0.57**) and were willing to continue using the tool in the future (**Mean = 4.60, SD = 0.67**). Overall satisfaction with IWCB was very high (**Mean = 4.45**). Additionally, students did not express major concerns about becoming dependent on the tool for generating ideas (**Mean = 4.15**), suggesting that IWCB was considered more as a supportive assistant than a replacement for independent thinking.

### Overall evaluation

The **overall average score** of the questionnaire was **4.28 (SD = 0.54)**, indicating strong positive perceptions of the effectiveness, motivational impact, and convenience of IWCB in English grammar practice. The relatively low SD values reflect a high level of consensus among students regarding the benefits and role of IWCB.

### *4.3 Interview*

Five students who used IWCB platform frequently volunteered to participate into the structured interview conducted online via Microsoft Teams, and they were coded as S1 to S5. Students were asked 4 questions so as to gather their experiences and perspectives towards using IWCB to practice IELTS writing.

The first question explored whether they feel motivated to practice writing IELTS essays thanks to the support of IWCB and discovered the factors that contribute to their motivation. All of the interviewees felt motivated and appreciated IWCB platform for various aspects it provided. S2 and S4 cited its time-flexible, need-based and personalized nature as what made them highly engaged. Furthermore, S3 reported that the instant, continuous feedback and offering scores across four IELTS criteria were the most useful features. In addition, S1 felt satisfied as she had a sense of progress when her initial scores given by IWCB gradually increased with every essay.

The second question inquired about the criterion which they improved the most among the four criteria thanks to the use of IWCB. All five students acknowledged that Grammar Range and Accuracy was the most significantly improved criterion. They reported that the IWCB identified and rectified grammatical errors as well as offering examples of complex sentence structures to help improve their writing band scores. In addition, S2 explained why the most significant improvement was seen in grammar: “*For grammar structures and rules, once I understand them, I can apply them to any writing task. However, the other criteria, especially vocabulary and ideas development depend a lot on the topic, so they’re harder to improve in a short time”.*

When it comes to the advantages of IWCB, all interviewees agreed that using IWCB for IELTS writing practice is an effective method, and the most evident advantage of IWCB is its **immediate feedback**. It gives scores, **clearly explains errors**, and **analyzes both strengths and weaknesses** across four assessment components, which helps learners make **instant improvements**. **S1** added a particularly insightful observation: "*IWCB gives corrections without changing my original ideas. This approach makes the feedback easier to understand and apply, and I am not worried about over-relying on the tool for ideas”.* S5 also also acknowledged the effectiveness of the learning method: “*When doing writing homework, I didn’t feel pressured to submit a perfect essay within a set time because the lecturer focused on assessing the revising process. What mattered most was showing that I could correct my mistakes and learn from the feedback*.”

The fourth question invited participants to share any concerns or aspects of IWCB they were dissatisfied with. Four students noted that using the free version of ChatGPT 3.5 limited them to about 10 rounds of essay revision per day, so they had to wait until the following day to continue. Only S5, who used the ChatGPT Plus version, was not affected by this limitation. S1 added an insightful observation: “*The feedback from IWCB is often quite general. Typically, IWCB provides examples of just a few errors for each criterion along with suggested corrections. Then I have to ask specific, focused questions to receive more detailed feedback. I think practicing essay revision with IWCB is an interactive, two-way process, so the effectiveness depends on whether the users know how to make further questions*”.

**5. Discussion**

After 12 weeks of using IWCB, the study results showed that the experimental group, which employed IWCB to practice IELTS writing, outperformed the control group. The post-test mean scores were also higher than the pre-test Mean scores, highlighting the positive influence of IWCB on students’ grammar competence.

The study's second research question explored students' perceptions of using IWCB for academic writing practice, especially in grammar skill, with 40 participants taking part in the survey. The findings revealed that most students had a positive perception of the effectiveness of IWCB in 4 IELTS writing components. Specifically, 90% of students selected “Strongly Agree” or “Agree” for items related to grammatical range and accuracy, followed by 82% for Task Achievement and Coherence and Cohesion, and 80% for Lexical Resource. This finding harmonizes with the results of several other studies (Liu et al., 2021; Song et al., 2023; Mahapatra, 2024). In addition, among the four components, IWCB was highly rated for grammar improvement while lexical resources had more students showing neutral and disagree with 20% altogether, indicating that vocabulary development still requires additional support from other sources. This result was confirmed by S2 from the interview and was explained by the rule-based nature of grammatical knowledge compared to the other components of IELTS writing which are more context-dependent.

When asked deeper into the improvement achieved in grammar accuracy and grammar range, the percentage of students who selected “Strongly Agree” and “Agree” for grammatical accuracy was 90% in comparison with grammatical range being lower at 77.5%. This indicates that IWCB was highly effective to support students in reinforcing their fundamental grammar knowledge such as verb tenses, articles, passive voice, comparatives, and punctuation. However, some students still had difficulty in using complex structures including relative clauses, compound sentences and conditional sentences. One possible explanation for this finding is that complex grammar structures require a deeper understanding of syntax, which often takes longer to master without lecturer’s guided practice or scaffolding. As the feedback provided by IWCB might not always offer step-by-step explanations or varied examples for these structures, additional teacher support might be necessary to help learners fully acquire and use these advanced grammar forms.

In terms of students’ motivation, nearly 90% of students found IWCB platform easy to use, expressed willingness to continue using it in the future, and were satisfied with its supportive role. Furthermore, nearly 80% felt motivated to practice IELTS writing thanks to the use of IWCB. These findings suggest that the use of IWCB contributed tremendously to boosting students’ motivation of learning IELTS writing, which resonates with the observations made by Song et al. (2023), who underscores the the positive influence of ChatGPT on Chinese EFL university students’ motivation and their heightened engagement in IELTS writing tasks.

When asked about their concerns regarding overreliance on IWCB for generating writing ideas, 75% of students reported that they did not rely heavily on the tool, indicating that IWCB effectively serves as a supportive tool rather than replacing learners’ creative thinking. This finding helps address concerns raised in previous studies by Barrot (2023), Yan, D. (2023), and Song et al. (2023), which warned that excessive dependence on AI tools could lead to plagiarism, reduced creativity, and diminished critical thinking skills. This revelation can be explained by the fact that the main function of IWCB is providing feedback and error correction rather than suggesting ideas or offering sample essays. In addition, the writing approach shifted the focus from perfection to progress, which assessed students’ progress through the revision process. Thanks to this approach, students did not feel pressurized to copy AI-produced content or the sample essay from Internet websites, and instead they were able to improve their writing skill through repeated practice.

The qualitative findings from the interview help reinforce, clarify and complement the quantitative results. The main advantages that students highlighted include time flexibiltiy, personalization, instant and continuous feedback, and the ability to offer scores across 4 IELTS criteria. Among these plus points of IWCB, the high appreciation of immediate and personalized feedback aligns with the revelations of some previous researchers including Barrot’s (2023), Han et al. (2023), and Mahapatra (2024) whose researches also underscored this feature of AI-supported writing tools tailored to students’ writing progress. Other than the perceived benefits which aligns with earlier research findings, students particularly valued the feature of providing IELTS scores for each writing component, which is typical of a tool specialized in assessing IELTS essays. S5 pointed out that this function helped him determine whether he was adequately prepared for taking the IELTS test or identify which aspects of his writing still needed improvement. S2 expressed her appreciation for IWCB’s ability to provide corrections without altering her original ideas, and S5 also gave credit for the writing practice approach which focused on essay revising process, which helped build their confidence in writing independently without relying on IWCB.

However, there were some downsides of using IWCB which were mentioned by the students, including limited number of feedback attempts when using the free version of ChatGPT, and a **restriction of** error correction feature **Chatbot can deliver in a single interaction**, requiring students to ask follow-up questions. In addition, unlike lecturer assessments where teachers can proactively recommend more advanced structures to enhance student writing, the IWCB shows a limitation in this regard in the initial round of feedback. Students need to take the initiative to ask follow-up questions targeting the specific aspects they wish to improve. This limitation of IWCB can be explained by the fact that chatbots are often optimized for speed which restricts the amount of content they can process at once (Gnewuch et al., 2022). In the first time of feedback, IWCB typically provides an overview of the strengths and weaknesses for each writing criterion, along with a few common errors and suggestions for improvement. If students require more detailed feedback, they need to continue interacting with the bot in subsequent exchanges.

**6. Conclusion**

This study investigated the effectiveness of the IELTS Writing Checker Bot (IWCB), a chatbot specializing in IELTS essays assessment, on the grammatical competence of EFL university students at Ho Chi Minh City University of Foreign Languages and Information Technology, and explored their perceptions of using this AI-powered writing tool. The findings offer significant contributions to the fields of EFL instruction and AI-assisted language learning.

Quantitative results revealed that there were considerable improvement in grammar performance of students who used IWCB, with post-test scores exceeding those of the control group and their own pre-test scores. Qualitative and survey findings further indicated a high level of student satisfaction with IWCB, especially in its ability to provide instant, personalized feedback and band scores aligned with the four IELTS writing components. Among these, the “grammatical range and accuracy” criterion received the highest appreciation, indicating IWCB’s effectiveness in reinforcing rule-based language features. Students also reported increased motivation and engagement in IELTS writing practice, while maintaining their autonomy and creativity, thereby addressing previous concerns about the overreliance on AI-supported tools. The study contributes valuable insights into how AI tools like IWCB can complement traditional instruction in assessing IELTS writing, which is meaningful in EFL contexts where access to individual feedback from teachers may be limited. However, some shortcomings of IWCB were also mentioned including the constrained feedback capacity with ChatGPT free version, specifically a lack of detailed error feedback and corrections in the first interaction, requiring students to take the initiative to ask for more targeted support.

Several limitations of this study should be acknowledged. First, the relatively small sample size and its confinement to a student group in a single university (HUFLIT) limit the generalizability of the findings. Second, the short duration of the intervention (12 weeks) may not fully measure and reflect long-term changes in writing proficiency. Therefore, further research should involve larger and more diverse samples across many institutions to enhance the generalizability of results as well as extending the duration of the intervention to examine long-term impacts of AI-supported writing tools. Beside, as this study focused specifically on the effectiveness of the IWCB in revising IELTS Writing Task 1, future research is recommended to investigate the impact of IWCB developed to revise IELTS Writing Task 2. Additionally, to address the current restrictions of IWCB tool, future work should focus on exploring the advanced versions of IWCB or similar tools which can be optimized to offer more comprehensive feedback in a single interaction.

In conclusion, this study confirms that AI tools like IWCB are highly effective in supporting EFL students with their academic writing. IWCB not only offers accessible and time-saving feedback but also helps assess IELTS essays accurately based on official scoring criteria. When used alongside classroom teaching and further improved, IWCB can play an important role in helping students develop their academic writing skills, especially in today’s technology-driven learning environment.

**References**

Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing student engagement: Harnessing “AIED”’s power in hybrid education - A review analysis. Education Sciences, 13(6), 632.

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745.

Golzar, J., Momenzadeh, S. E., & Miri, M. A. (2022). Afghan English teachers’ and students’ perceptions of formative assess- ment: A comparative analysis. *Cogent Education, 9*(1), 2107297.

Gnewuch, U., Morana, S., Adam, M. T., & Maedche, A. (2022). Opposing effects of response time in human–chatbot interaction: The moderating role of prior experience. *Business & Information Systems Engineering*, *64*(6), 773-791.

Guo, Q., Feng, R., & Hua, Y. (2022). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing?. *Computer Assisted Language Learning*, *35*(9), 2312-2331.

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., ... & Oh, A. (2023, July). RECIPE: How to integrate ChatGPT into EFL writing education. In *Proceedings of the tenth ACM conference on learning@ scale* (pp. 416-420).

Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. Journal of Physics: Conference Series, 1193(1), 012020.

Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions. Journal of Computing in Higher Education, 35, 40–68.

Hyland, K. (2003). Writing and teaching writing. *Second language writing*, *1*(2), 1-30.

Johri, P., Khatri, S. K., Al-Taani, A. T., Sabharwal, M., Suvanov, S., & Kumar, A. (2021). Natural language processing: History, evolution, application, and future work. In Proceedings of the 3rd International Conference on Computing Informatics and Networks (ICCIN 2020) (pp. 365–375).

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability, 15(12), 12451.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal*, *54*(2), 537-550.

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. Sustainability, 15(7), 5614.

Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2023). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, *31*(9), 5614-5632.

Mahapatra, S. (2024). Impact of ChatGPT on ESL students’ academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, *11*(1), 9.

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.

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.

Skrebeca, J., Kalniete, P., Goldbergs, J., Pitkevica, L., Tihomirova, D., & Romanovs, A. (2021). Modern development trends of chatbots using artificial intelligence (AI). In 2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS) (pp. 1–6).

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.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752.

Tai, A. M. Y., Meyer, M., Varidel, M., Prodan, A., Vogel, M., Iorfino, F., & Krausz, R. M. (2023). Exploring the potential and limita- tions of ChatGPT for academic peer-reviewed writing: Addressing linguistic injustice and ethical concerns. *Journal of Academic Language and Learning, 17*(1), T16–T30.

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943-13967.

Zhao, X. (2022). Leveraging artificial intelligence (AI) technology for English writing: introducing Wordtune as a digital writing assistant for EFL writers. RELC J. 89:10940.

Zhu, M., Liu, O. L., & Lee, H. S. (2020). The effect of automated feedback on revision behavior and learning gains in forma- tive assessment of scientific argument writing. *Computers & Education, 143*, 103668.

**Bionote**

Nguyen Thi Huyen Tran is currently an EFL lecturer of the Faculty of Foreign Languages at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT). She has been teaching English for more than 15 years. Her interested research focuses on Teaching English to Speakers of Other Languages (TESOL) and AI-Assisted Language Learning. Email: trannth1@huflit.edu.vn

**APENDIX – THE PROMPT FOR IELTS WRITING CHECKER BOT**

This is a standard prompt for Academic writing Task-1 Assessment only.

It requires the input with 3 factors:

1- Task 1 questions (Academic Style)

2- Task 1 Visual picture

3- Task 1 report

And if one of 3 factors is missing, please ask for the missing one before the prompt below is active.

**The active prompt in non-Public Examiner Descriptors:**

Please assume the role of a certified IELTS Examiner with 20 years of experience at the British Council, specializing in Academic IELTS Writing assessment and test design. You are trained to assess Academic writing Task-1 report by using the official Examiner Band Descriptors (non-public version), and are responsible for assigning Band scores accurately according to standardised assessment protocols.

Please evaluate the student’s Task 1 response based on the following input:

**Input**

a) Task 1 question: [Insert full Academic Task 1 prompt]

b) Visual (chart/graph/map/process): [Insert description or image]

c) Student’s written report: [Insert full response]

**Scoring Criteria:**

Assess the report using the four official IELTS Writing Band Descriptors, and assign a Band Score (1–9) for each:

1. Task Achievement (TA)

Evaluate if the response:

* select and group key features of the information.
* provide sufficient detailto illustrate these features.
* report the information, figures and trends accurately.

Compare or contrast the information by adequately highlighting the identifiable trends, principal changes or differences in the data and other inputs (rather than mechanical description reporting detail).

Present the response in an appropriate format

Avoids irrelevant detail or personal opinion.

Band Score: \_\_

2. Coherence and Cohesion (CC)

Assess whether the report:

* has the coherence of the response via the logical organisation of information and/or ideas
* has logical sequencing of ideas and/or information within and across paragraphs.
* has appropriate and flexibleuse of cohesive devices (discourse markers, reference and substitution,..)to form logical relationship between ideas.
* has the appropriate use of paragraphing for topic organisation and presentation.

Band Score: \_\_

3. Lexical Resource (LR)

Examine if the writer:

* uses a wide and accurate range of vocabulary related to data description.
* avoids repetition and uses appropriate synonyms/paraphrasing.
* uses collocations and sophisticated phrasing.
* controls spelling and word formation.

Band Score: \_\_

4. Grammatical Range and Accuracy (GRA)

Determine if the writing:

* demonstrates a wide range of sentence structures (simple, complex, compound).
* has high grammatical accuracy with only occasional slips.
* has correct use of grammar such as tense, passive voice, comparisons, quantifiers, and punctuation, etc

Band Score: \_\_

**For Each Criterion, provide a structured assessment including:**

Error:

* Quote examples of grammatical, lexical, cohesion, or structural mistakes.
* Highlight misused transitions, incorrect sentence forms, punctuation, or spelling issues.

Improvement suggestion:

* Recommend specific corrections to improve clarity, conciseness, and accuracy.
* Suggest better sentence structures or vocabulary replacements.

Explanation:

* Briefly explain how the error affects the score in this criterion.
* Clarify how the correction would raise performance to a higher band.

Examples (Before → After):

Format your examples like this:

Incorrect: The number increased dramatical in 2010.

Corrected: The number increased dramatically in 2010.

Note: Students are encouraged to self-revise based on feedback and rewrite the essay, so DO NOT provide model paragraphs or sample essays.

Overall Band Score:

* Provide a calculated overall band score (average of the four criteria).
* Add 1–4 lines of final advice on how the student can reach a Band 7.0+ performance.