

Effectiveness of ChatGPT Feedback on Low-Level English Students' Writing Proficiency and Their Perceptions: An Experimental Study

Vu Quynh Hoa^{1*}, Nguyen Kim Phuc¹

¹ Faculty of Foreign Languages, Hanoi University of Science and Technology, Vietnam

*Corresponding author's email: hoa.vuquynh@hust.edu.vn

 <https://orcid.org/0009-0008-4785-1132>

 <https://doi.org/10.54855/ijaile.26324>

® Copyright (c) 2026 Vu Quynh Hoa, Nguyen Kim Phuc

Received: 22/06/2026

Revision: 16/08/2026

Accepted: 20/08/2026

Online: 31/08/2026

ABSTRACT

Keywords:

ChatGPT, Writing skills, Low-level students

The purpose of this study is to explore the effectiveness of ChatGPT's feedback on the writing of non-English major university students and their perceptions of this implementation. Participants in this study had a low level of English (A2). Researchers employed an experimental research design to achieve the research objectives. There were 42 students in total, including 21 students in the experimental group (EG) and 21 in the control group (CG). EG received ChatGPT feedback after writing draft 1, whereas CG received peer feedback and general instructions from teachers. Data were collected through pre- and post-tests, and semi-structured interviews. Paired sample t-tests revealed a significant improvement in mean scores between pre-and post-tests of EG ($p=.00<.05$), in contrast to CG's. While students complimented ChatGPT on its immediacy, dialogic, and personalized responses, the tool's prompting difficulties and blind acceptance were their major concerns. Findings suggest the need for clear AI literacy training for students and a 'human-in-the-loop' approach in which ChatGPT feedback incorporates teachers' guidance.

Introduction

In recent years, the use of ChatGPT to provide formative feedback in writing instruction has gained considerable attention from researchers. Studies show its potential in improving learners' writing proficiency (Abdi et al., 2024; Lai & Trinh, 2025; Zhang & Chen, 2026). Specifically, for learners, ChatGPT provides instant, detailed, and personalized feedback; meanwhile, for teachers, ChatGPT, like other AI generative tools, is regarded as a productive teaching assistant that helps teachers give learners timely feedback and reduces workload in large classrooms (Ardito, 2024).

Since its public release, there has been plenty of research studying the use of ChatGPT in education in general and teaching English writing in particular. When conducting research into ChatGPT's effectiveness in giving writing feedback, the researchers focused on recent, reliable, and insightful studies with similar objectives. In general, studies confirm ChatGPT's potential in giving students automated feedback.

Across these studies, ChatGPT has been associated with measurable gains in writing scores, grammar accuracy, thesis development, idea organization, and expression (Meniado et al., 2024; Poláková & Ivenz, 2024), particularly when combined with teacher feedback (Asadi et al., 2025) or applied to specific components of writing such as grammar, coherence, and vocabulary (Wang et al., 2023; Chan & Hu, 2023). At the same time, comparative and qualitative studies caution that ChatGPT still falls short of human feedback in analytical depth, personalization, and context-sensitivity (Steiss et al., 2024; Ljubojević, 2024). A more recent Vietnamese study further found that ChatGPT tends to assess student writing differently from teachers, often assigning notably higher scores across rubric criteria, with the greatest discrepancy occurring in content evaluation (Ho et al., 2026). This pattern is corroborated by a broader review of ChatGPT's role in written corrective feedback, which similarly found that while the tool offers benefits such as immediacy and clarity, its effectiveness for advanced writing skills, rhetorical elements, and cultural nuance remains limited compared to teacher feedback (Nguyen, 2026).

Despite promising results, a closer look at the studies reviewed above reveals a gap: none specifically targeted low-level English language learners, and their focus remained mainly on writing tasks rather than free writing. Consequently, how ChatGPT's feedback supports low-level learners in free writing tasks remains largely unexplored. The research team conducted this study to provide more empirical evidence for the role of ChatGPT in giving learners' writing formative feedback, simultaneously clarifying its role in supporting sustainable and autonomous learning skills.

Literature Review

Formative feedback in writing

Formative feedback in writing is generally understood as comments and suggestions on a draft, aimed at helping writers revise it to improve their products (Texas A&M University, 2026). Formative feedback plays a key role in improving learners' writing ability. According to Steiss et al. (2024), formative feedback in the writing process helps students understand what constitutes a quality piece of writing and how to improve different aspects of their writing. This, hence, helps plan suitable actions and effective learning strategies (Graham et al., 2011; Panadero & Jonsson, 2013). Therefore, formative feedback is not only a supporting tool for writing skills but also nurtures independent learning skills and critical thinking for learners (Lam, 2018; Mahapatra, 2024). So far, two types of formative feedback have been recognized. They are (1) traditional formative feedback, which is completely given by human beings (teachers/ peers), and (2) automated feedback, which is given by machines.

However, in practice, conventional formative feedback, which requires personalization for learners, takes considerable time and effort, leading many teachers to hesitate or refuse to do the task. This calls for solutions that reduce teachers' workload while still ensuring quality formative feedback for learners. Effective formative feedback, according to Steiss et al. (2024), includes clarity, precision, and a focus on the important points in the writing, with supportive tone. Hattie & Timperley (2007) claimed that students' motivation might increase, thereby improving writing skills, thanks to supportive, non-directive feedback in a supportive tone. These requirements were somewhat addressed by automated writing feedback, as students could receive fast, constructive feedback, which helped learners improve their writing results significantly (Fleckenstein et al., 2023).

Formative automated writing feedback

However, it was noted that early automated writing feedback tools had limitations, including low precision and the need for complex data training (Shermis & Wilson, 2024). In this context, the development of AI tools such as ChatGPT brings new opportunities because it could resolve the limitations above (Kurt and Kurt, 2024). For example, according to Steiss et al. (2024), ChatGPT could provide more flexible, precise, and personalized responses. Besides, Asadi et al. (2025) claimed that ChatGPT could also stimulate comprehensive development in language learning and students' motivation.

Research by Poláková and Ivenz (2024) shows that writing skills, distinct from other English skills, require formative feedback and repeated practice. Therefore, ChatGPT, with its instant and continuous responses, has proven its value in promoting learners' progress. Besides, Poláková & Ivenz (2024) claimed that formative feedback is commonly used in self- and peer-assessment, thereby promoting reflection, self-regulation, and self-correction in students' writing (Lam, 2018). With this idea, automated feedback has become an important part of the automation and upgrade of formative feedback. Studies by Rahman et al. (2022) show that students respond positively to automated feedback for identifying grammatical errors and improving writing skills. Automated feedback not only reduces teachers' burden (Wilson et al., 2022) but also contributes to learners' self-regulation (Grimes & Warschauer, 2010).

However, it is noted that ChatGPT should be used as a supporting tool, complementing teachers' feedback rather than completely replacing feedback sources (Chun et al., 2021; Zhang & Hyland, 2018). One notable reason was ChatGPT's misidentification of students' mistakes due to its preconfigured prediction algorithm (Hasanah & Safitri, 2024). The combination of AI-automated feedback and insightful feedback from teachers promises to create a flexible learning environment that develops writing skills in a sustainable and more rounded way.

ChatGPT prompting techniques

Designing an effective prompt for ChatGPT to provide formative writing feedback to low-level English students is key to exploiting this AI tool's potential. This group of students often struggles with presenting ideas, organizing essays, and applying correct grammar (Firkins et al., 2007; Aldabbus & Almansouri, 2022). Therefore, formative feedback needs to be clear, simple, and step-by-step. To achieve this, prompting techniques introduced by Brookman (2025, as cited in Stillman, 2025), which considered CLEAR (Lo, 2023), IDEA (Google for Educators, 2024), and the zero-shot prompting technique (Chae & Davidson, 2025), have been taken into account for the purpose of suggesting well-made prompts for students.

CLEAR (Lo, 2023) emphasizes 'concise, logical, explicit, adaptive and reflective' in prompts. Meanwhile, IDEA (Google for Educators, 2024) adds personalized factors for prompts through 'personal role, aim, recipients, theme, and expected structure'. This is particularly crucial in the context of learning to write, as feedback should be precise and suitable for learners' level. In addition, zero-shot, which means providing only questions without training the tool or sample examples, was chosen to introduce to students in this research to simplify the interaction process. Finally, the prompting technique outlined by Brookman (2025, as cited in Stillman, 2025) could be considered a practicalization of the theoretical frameworks discussed, including CLEAR and IDEA. Therefore, essential factors in the prompt include: (1) state your goal, (2) specify your preferred format, (3) warnings and guardrails, and (4) context dump. Beyond these structured frameworks, Woo et al. (2023) observed that non-technical users, including EFL students, often approach prompt writing through a trial-and-error process, gradually refining their prompts based on the chatbot's responses.

To sum up, formative feedback has long been recognized as a key factor in improving EFL learners' writing skills. However, formative feedback that meets essential criteria, including clarity, precision, and personalization, is time-consuming and effort-intensive for teachers. While the development of AI tools such as ChatGPT has shown potential in reducing teachers' workload by providing instant and personalized feedback, some scholars emphasize that ChatGPT should serve as a supporting tool for teachers in giving writing feedback rather than completely replacing them. At the same time, finding a suitable prompt in this context for low-level English students is essential for assessing the effectiveness of ChatGPT's feedback in enhancing students' writing skills. With this existing literature, we aim to explore (1) the effectiveness of ChatGPT feedback on low English level students' writing ability compared to traditional ways of giving feedback in our teaching context; and (2) students' attitudes towards the experience of working with ChatGPT in learning writing.

Research Questions

The study aims to address two research questions:

1. How did ChatGPT feedback affect low-level English students' writing ability compared to traditional ways of giving feedback?
2. What were students' attitudes towards ChatGPT feedback applied in their English writing lessons?

Methods

Pedagogical Setting & Participants

The research was conducted in two General English classes in a science and technology university that one of the two researchers taught during the second term of the 2024-2025 school year. In total, students studied 6 units, including reading, listening, communication skills, and writing lessons. Therefore, all 6 writing lessons of the 6 units in EG received intervention. Before the intervention, all participants were informed about the intervention by the teacher, who was also the researcher. Each of them then received an informed consent form. After receiving confirmation of participation, the researchers proceeded to the next stages of the research.

All students specialized in technical fields; they studied English as a mandatory subject in the first year of university. A convenience sampling method was chosen; students in both classes were at the A2 level in the CEFR, and class sizes were roughly the same. EG had 24 students, and CG had 23 students; however, two students in EG were excluded because of their absence in many lessons, and one student in CG was not suitable for the test because that student's English level was B2 (he was in class because he did not show up in the replacement test). Therefore, 21 students in each group participated in the study.

Design of the Study

Figure 1.

Research design

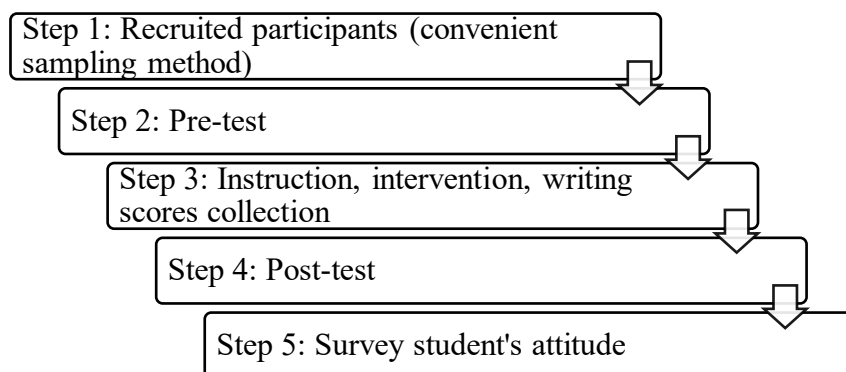


Figure 1 maps the research design. To achieve the research aims, the researchers first selected participants using convenience sampling, as they were the two most suitable and accessible groups. Then, in the second step, students were asked to take pretests to assess their initial writing proficiency. In the next step, the teacher gave EG students instructions on using ChatGPT feedback for their writing compositions, then began the intervention and collected writing scores. The details of the intervention process are presented in the next paragraph. In the fourth step, researchers administered post-tests to assess changes after the intervention, and, finally, we surveyed students' attitudes towards the intervention through semi-structured interviews.

The intervention was conducted as follows. In writing lessons, the teacher followed a step-by-step process: pre-writing, teaching students to write first drafts, providing feedback, and then revising drafts. For EG, after writing the first drafts and receiving the teacher's instruction, students used ChatGPT to get feedback; they then studied the feedback and wrote revised drafts based on ChatGPT's comments. Meanwhile, CG learned using traditional methods, which means they received peer feedback and learned from teachers' feedback for some random compositions in class. In total, there were six writing lessons.

After the intervention, five students from EG were selected to participate in the semi-structured interviews based on their availability and their writing scores in both the pre-test and post-test. We also aimed to ensure diversity among the participants by selecting individuals with a wide range of scores - from the lowest to the highest. Specifically:

Table 1.

Information of participants taking part in semi-structured interviews

Student code	Gender	Pre-test score	Post-test score
S1	Female	Lowest - 39	Lowest - 55
S2	Male	Low - 42	Low - 57
S3	Male	Medium - 58	Medium - 71
S4	Female	Medium - 58	Medium - 68
S5	Male	Highest - 60	Highest - 82

Data collection

In this research, we collected both quantitative data and qualitative data. Quantitative data were students' writing scores, which were gathered through pre-tests, post-tests, and three randomly

revised writing drafts in three lessons. Qualitative data, which were students' accounts of their attitudes towards the intervention experience, were gathered through semi-structured interviews. To collect data, an interview guide was designed based on open-ended questions covering two broad aspects: (1) students' attitudes towards the experience using ChatGPT to get formative feedback regarding its usefulness and ease of use, and (2) students' experience with generating prompts for ChatGPT feedback, including the language and procedures they used, the difficulties they encountered, and their suggestions for applying ChatGPT feedback effectively in writing lessons.

Data analysis

The quantitative data were students' writing scores; the researchers used Jacobs et al.'s (1981) ESL composition profile as an official, reliable scoring rubric to evaluate students' writing in both CG and EG. This was because the scoring rubric was claimed to provide a comprehensible evaluation of writing products, especially for teaching and scoring ESL learners. (Setyowati et al., 2020). Scores of pre-tests, post-tests, and three random pieces of writing were then compared by an independent t-Test. Regarding qualitative data, students' attitudes and responses from five students across three levels of performance in EG were transcribed and analyzed for themes.

Results and Findings

Research question 1: How did ChatGPT feedback affect low English level students' writing ability compared to traditional ways of giving feedback?

To evaluate the effectiveness of the intervention, the mean, SD, and p-value from an independent-samples t-test were used to compare the average scores of the Control Group (CG) and Experimental Group (EG) at different time points.

Table 2.

Results of independent t-tests (pre-test)

Pre-test	Mean	Std. Deviation	p-value
CG (n = 21)	50.76	9.48	> 0.05
EG (n = 21)	51.90	7.02	

Before the intervention, pre-tests showed little statistical difference between the two groups. The mean score of CG was 50.76 (SD = 9.48), while EG's was 51.90 (SD = 7.02). $p > 0.05$ indicates no statistically significant difference between the two groups. This confirms the similarity in writing ability between the two groups at the beginning of the study, paving the way for evaluating the effectiveness of the intervention.

Figure 2.*Results of three pieces of writing (revised drafts)*

The researchers selected three writing lessons to compare students' writing progress during the intervention. In writing 1, EG reached 70.9 for mean score (SD = 9.75), as opposed to CG's mean score of 55.61 (SD = 9.42), p -value < 0.05 . This difference indicates statistical significance. In the second and third writings, the trend continued, with $p < 0.05$, affirming the positive impacts of ChatGPT's feedback on students' writing. In the third piece of writing, although EG's mean score was almost the same as the previous writing, the difference with CG was stable, with statistical significance. This indicates that EG tended to reach a 'threshold point', while CG continued to show a slow but steady increase.

Table 3.*Results of independent t-tests (post-test)*

Post-test	Mean	Std. Deviation	p-value
CG (n = 21)	60.81	9.08	< 0.05
EG (n = 21)	69.85	8.66	

Table 3 shows the post-test results for two groups, confirming an overall trend. The mean score of CG was 60.81 (SD = 9.08), and EG's mean score was 69.85 (SD = 8.66), $p < 0.05$. The difference is statistically significant, indicating that the intervention considerably improved students' writing ability compared to traditional teaching methods.

Table 4.*Results of independent t-tests regarding improvement in mean scores between pre-and post-test*

Group	Pre-Test	Post-Test	p-value
CG (n = 21)	50.76 ± 9.48	60.80 ± 9.08	< 0.05
EG (n = 21)	51.90 ± 7.02	69.85 ± 8.66	< 0.05

Table 4 reveals a significant improvement in mean scores between pre-and post-tests of EG ($p = .00 < .05$), in contrast to CG's.

Research question 2: What were students' attitudes towards ChatGPT feedback applied in their English writing lessons?

All participants generally shared positive thoughts towards ChatGPT feedback, claiming that it was prompt and detailed. As S5 explained,

Unlike humans, the software doesn't care whether its responses hurt or offend users. It simply points out the problems directly and specifically. That's especially helpful for students with lower English proficiency because they have many knowledge gaps and need clear feedback to improve their writing.

All five interviewees also affirmed that ChatGPT feedback helped them fix grammar, vocabulary, and spelling errors. In particular, this tool helped improve students' "use of verb tenses" (S4), provide grammatical knowledge that students had misunderstood or had not yet known (S2), and "suggest more accurate word choices based on the given contexts" (S3). Furthermore, ChatGPT feedback was said to enhance students' content organization. According to S2 and S4, by helping them restructure their writing outlines, this AI tool not only turned the sketchy original drafts into more logical ones but also improved their thought flow. For S5, the most significant strength of ChatGPT is its ability to present information logically and coherently. This allows students with low English levels to imitate it to strengthen their reasoning.

However, the tool still contains certain weaknesses. As stated by S3 and S4, although ChatGPT could often provide specific and prompt evaluation of writing quality, students themselves had to consider whether the issues raised by this software were valid. For instance, sometimes ChatGPT "falsely identified the verb tense error" as it misunderstood the context (S4), or "sometimes it would provide too general feedback that I could not identify what the error was" (S2). This could be challenging for students with low English proficiency, since they tend to make many errors and may not recognize all of them. Besides, this AI tool could fail to answer students' questions correctly and even provide confusing information in its responses (S5). Notably, when S1 asked ChatGPT to evaluate her use of punctuation and capitalization, the software completely ignored the request, focusing solely on analyzing her grammar and vocabulary errors. This may hinder students from developing comprehensive writing skills, as the software overlooks important criteria they must follow.

Students' prompt-generating experience

a. Language use for generating prompts

Most participants utilized Vietnamese during their conversations with ChatGPT. Explaining this, S3 shared that writing prompts in the mother tongue helped him express requests more clearly, and ChatGPT could also rely on that to respond more appropriately. For S1, her current low level of English made it impossible for her to express her thoughts in that language. Therefore, she chose to use Vietnamese to communicate with the tool instead. She also argued that

if I thoroughly inquire and know how to harness AI data effectively, then even if I interact with ChatGPT in Vietnamese, I can still collect useful knowledge and thereby enhancing my English writing proficiency.

Among all participants, only S5 used a mixture of Vietnamese and English to communicate with ChatGPT. However, he stated that when he mixed both languages in particular prompts, ChatGPT tended to respond in English only. His solution was to keep reading the AI feedback.

If he did not understand anything, he would ask the software to simplify or translate such information into Vietnamese.

b. Prompt generating procedure

Although the general prompt-generating procedure had been taught in class, participants affirmed that they gradually adjusted and tailored it to make it more convenient for them.

Differences in the prompt-generating procedure among participants can be illustrated in the following step-by-step workflows:

Figure 3.

The prompt-generating procedure of S3 and S4

Step 1: Provide basic requirements

- Provide feedback for my writing. The word limit required is ... words
- The writing question is ...
- Provide feedback on grammar, vocabulary and content of the writing
- What do I need to improve?
- I am at A2 level, so please respond using A2-level English.

Step 2: Provide the full writing

Step 3: Receive AI feedback. Use follow-up prompts for deeper insights

- Clarify this error. Why is it an error and how do I fix it?
- You noted this part as problematic. How is it so?
- How to further develop the content of this part?

Change this word into a more appropriate one

Figure 4

The prompt-generating procedure of S1 and S2

Step 1: Provide the full writing immediately

Step 2: Notify ChatGPT that I am at A2 level, so it should respond accordingly

Step 3: Request an overall score

Step 4: Ask ChatGPT to provide more specific feedback related to grammar, vocabulary and content of the writing

Figure 5

The prompt-generating procedure of S5

Step 1: Self-evaluation- Review the writing carefully, take notes on identified weaknesses

Step 2: Provide basic requirements, also ask ChatGPT to focus on weaknesses that were identified

Step 3: Provide the full writing

Step 4: Receive AI feedback. Use follow-up prompts for deeper insights, especially focusing on targeted areas to improve writing skill

S5 stated,

Writing ability comes from both talent and practice. After practicing writing English for a while, I believe I should only focus on my weaknesses and use ChatGPT to improve them in the most effective way. This not only saves time but also enhances self-learning efficiency. If I simply share the whole writing and ask ChatGPT to correct it, the results I receive would be too vague- the software provides too much information, making it difficult to filter out the most relevant and necessary points to me.

c. Challenges in prompt generation

S1 complained that ChatGPT often automatically rewrote her entire piece of writing, even though she only needed it to provide feedback and fix her errors. Besides, ChatGPT sometimes did not understand prompts provided by students. Consequently, its responses did not meet their expectations (S2, S3 and S4).

Thanks to the considerable time spent interacting with ChatGPT, all participants gained experience dealing with such issues. For instance, S2 would simplify his prompts using simpler lexis and shortened expressions. S3 gave ChatGPT an example to illustrate his point clearly before asking the tool to respond.

With S4, her problem was that ChatGPT sometimes forgot what her English level was and thereby provided feedback with much more complex language: *“Sometimes, ChatGPT provides very advanced sentences with lots of words and structures that I do not understand”* (S4). To solve such a problem, she had to keep reminding the software of her English proficiency.

Students’ suggestions to apply ChatGPT feedback for effective writing

S2 believed ChatGPT should be used during the Pre-writing stage to aid in brainstorming and structuring content, especially since he struggled to organize ideas due to limited writing experience. He suggested that if ChatGPT generated initial content in Vietnamese, it could help him apply grammar and vocabulary more effectively. In contrast, S4 and S5 preferred to use ChatGPT during the Drafting and Revision-Feedback stages, after students had already developed their writing independently. S5 strongly opposed its use in Pre-writing, arguing that early reliance on AI could hinder students’ original thinking.

All participants planned to continue using ChatGPT feedback for their writing practice but emphasized the need to receive evaluations from both the software and their teachers. S2 shared that while teachers provide tailored feedback based on each student’s proficiency and learning needs, ChatGPT could help support large classes, especially when there is not enough in-class time for feedback. S5 agreed, noting that while ChatGPT’s polished style can feel mechanical and difficult to understand, teachers suggest using accessible, natural language for weaker students. For S4, since this AI tool sometimes provides too general feedback, students must actively seek additional information to gain insights into certain aspects of their writing, for example, by crafting specific prompts to receive responses tailored to their needs. In contrast, teachers provide comprehensive feedback on all aspects of students’ writing, allowing them to simply listen, reflect, and improve without spending extra time searching for information. This makes the learning process more comfortable and efficient. To save time and ensure that teachers can still review each student’s work, S2 and S3 suggested that students should first share their writing with ChatGPT for initial feedback. Afterward, they can share such AI-generated comments with their teachers, allowing the teachers to evaluate their accuracy and provide additional feedback to further refine the writing.

Discussion

Quantitative data indicate that students who received ChatGPT's feedback showed an apparent improvement in writing ability compared to those who didn't. This result is similar to previous studies about automated writing feedback tools in teaching English writing. For instance, Mahapatra (2024) reported that ChatGPT, as a generative feedback tool, had positive effects on ESL students' academic writing ability. Systematic reviews also reported that automated writing feedback can provide constructive feedback, leading to significant improvements in students' writing (Fleckenstein et al., 2023). These findings affirm that the use of AI-based technology such as ChatGPT can help students enhance their writing skills by providing timely and personalized responses. Research findings are also similar to surveys on ChatGPT's effectiveness in language education. For example, Kurt and Kurt (2024) found that students accepted ChatGPT for its ease of use and useful responses. Simultaneously, Kurt and Kurt's (2024) study warned that ChatGPT can only solve linguistic problems on the surface and sometimes provides inconsistent answers; therefore, it is impossible to completely replace teachers' or peers' feedback with ChatGPT's. Whether participants in this research provided similar opinions will be presented in the discussion for the second research question.

Regarding the effectiveness of ChatGPT feedback, the qualitative findings indicate that this tool enhanced students' writing skills, particularly in grammar, vocabulary, and content organization. This is similar to Ljubojević (2024), in which ChatGPT was found to help students improve their understanding of the structure and genre of their writing, generate ideas, and enhance their English language skills.

However, providing inaccurate feedback is one of the tool's typical weaknesses. For our study, participants realized that ChatGPT sometimes incorrectly detected grammatical errors. Similarly, as stated in Hasanah & Safitri (2024), ChatGPT's misidentification of the student's mistake suggests that, rather than understanding the given text, this software relies on a predictive algorithm to generate its response. Notably, this automated writing evaluation tool could be overly general and wordy and may at times confuse those receiving feedback (Grimes & Warschauer, 2010; Shermis & Wilson, 2024). This is also consistent with our interviewees' opinions. According to Grimes & Warschauer (2010), inaccurate or confusing feedback can discourage students, especially novice writers, from engaging in revision. To address this problem, Hasanah & Safitri (2024) suggested that teachers remind students that AI-generated feedback is not always accurate, and that students should carefully consider its accuracy and usefulness before integrating it into their writing. Since students in our study had low English proficiency, teachers could scaffold their reflections.

Based on our findings, 2 students with the lowest pre- and post-test scores (S1, S2) used a different prompt-generating procedure than the 3 other students with higher pre- and post-test scores (S3, S4, S5). In particular, S1 and S2 initially provided ChatGPT with their writing and then generated additional prompts to elicit ChatGPT's feedback. This procedure reflects *trial-and-error prompting*, in which students adjusted their prompts in real time based on the responses received. They learned through repeated attempts which types of prompts best elicited the feedback they needed (Woo et al., 2023). Meanwhile, S3, S4, and S5 chose to provide ChatGPT with the essential requirements first, then shared their writing with the tool. The strategy applied here was *chain prompting* - a structured prompting strategy where users break down their requests into sequential steps, ensuring that AI responses are progressively refined and contextually relevant (Wei et al., 2022). Although there have been no studies directly proving the relationship between prompt-generating strategies and students' English proficiency, the differences observed in prompt-generating procedures suggest that higher-

proficiency students tend to engage in more structured approaches like chain prompting, as they have a clearer understanding of their needs and expectations, whereas lower-proficiency students often engage in trial-and-error prompting – adjusting their requests based on AI responses. This implies that, in addition to language ability, to apply AI feedback effectively in English writing, students need to develop strategic thinking and cognitive flexibility when crafting prompts. Teachers may consider providing students with further guidance and practice on these elements, along with structured prompting techniques, so they can better control AI response quality, especially since prompting affects the type of feedback generated (Steiss et al., 2024).

Although S5 appreciates ChatGPT's straightforward approach to feedback, which offers thorough comments on low-level students' writing issues, Hattie & Timperley (2007) suggested that students may feel more motivated and improve their writing skills more effectively when receiving feedback in a *supportive, non-directive* tone – similarly to what is typically provided by human teachers. Although this has not been verified in the study, apart from S5, no other participants reported that the straightforward versus supportive tone of feedback affected their writing skills; students generally expressed a strong preference for receiving feedback from teachers alongside ChatGPT. Some even noted that teachers provide deeper, more tailored insights, making their feedback more suitable for individual students than ChatGPT's. This is similar to Steiss et al. (2024), who indicated that human evaluators, particularly those who are well-trained, compensated, and have sufficient time, deliver higher-quality feedback in terms of clarity in guiding improvements, accuracy, and prioritization of essential aspects.

In our study, S5, a higher-proficiency student, opposed applying ChatGPT at the Prewriting stage, emphasizing independent idea generation and critical thinking. This is consistent with Hasanah & Safitri (2024), who stated that dependence on technology from the outset could diminish creativity and critical thinking. As students increasingly use AI-powered tools for idea generation and writing, they may rely less on their own cognitive processes. Over time, this reliance could weaken their ability to think independently and innovate. However, S2 – a lower-proficiency student – was supported by ChatGPT during the Prewriting stage because of his limited writing experience and strong desire to learn from AI-generated sample content. These contrasting experiences suggest that writing proficiency significantly shapes both the approach and extent to which students utilize ChatGPT feedback. It also influences whether they use AI as a refinement tool or rely on it for foundational guidance.

Participants' suggestion to combine feedback from both teachers and ChatGPT also aligns with Steiss et al. (2024), who recommend integrating AI-generated feedback for first drafts while reserving human evaluation for later or final revisions. This approach ensures that the human perspective remains central and allows instructors to address any shortcomings in AI feedback, such as inaccuracies or failure to emphasize critical areas for improvement.

Conclusion

This research has provided empirical evidence of the effectiveness of ChatGPT in providing formative writing feedback to low-level English learners. Quantitative data indicate that students who received ChatGPT's feedback significantly improved their writing skills compared to those in the CG. Simultaneously, qualitative data indicated that this tool could help learners refine their grammar, vocabulary, and organization. Additionally, they complimented this tool's immediacy, dialogic nature, and personalized responses. Regarding variations in prompting techniques (e.g., trial-and-error prompting and chain prompting), the study revealed

that higher-level students tended to approach ChatGPT more systematically and with a clearer goal, whereas low-level students were inclined to adjust their prompts based on the AI tool's responses.

However, the study also spotted some notable drawbacks of this means of feedback. Firstly, feedback was sometimes inaccurate or overly general, causing confusion among learners. Furthermore, the reliance on ChatGPT in brainstorming could reduce learners' critical thinking and creativity. Even though some learners appreciated ChatGPT's generated feedback, they still expected to combine teacher and ChatGPT feedback to ensure personalization and depth of comment. This highlights the necessity of teachers' role in supervising, supporting, and instructing the use of the AI tool.

This study has certain limitations. First, the sample size was small, and the study was limited to low-proficiency learners; therefore, the findings may not generalize to other types of learners. Secondly, there was no control over learners' interactions with ChatGPT while receiving feedback, making the effectiveness dependent on several factors, such as prompting ability, understanding of ChatGPT's responses, and learners' motivation.

The study suggests that future research should expand its scope and compare proficiency levels. Additionally, there should be more insightful studies on prompting strategies and the correlation between learners' understanding of AI feedback and their ability to optimize responses; all of these aim to enhance the implementation of ChatGPT and other generative AI tools in writing instruction.

References

- Abdi, M. M., Ananta, B. D. B., & Nurbatra, L. H. (2024). Exploring the use ChatGPT in improving students' paragraph composition skills. *Focus on ELT Journal*, 6 (4), 17-36. <https://doi.org/10.14744/felt.6.4.2>
- Aldabbus, S., & Almansouri, E. (2022). Academic writing difficulties encountered by University EFL Learners. *British Journal of English Linguistics*, 10(3), 1–11. <https://doi.org/10.37745/bjel.2013/vol10n3111>
- Ardito, C. G. (2024). Generative AI detection in higher education assessments. *New Dir Teach Learn*, n/a(n/a). <https://doi.org/10.1002/tl.20624>
- 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>
- Chan, C., & 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), 1–18. <https://doi.org/10.1186/s41239-023-00411-8>.
- Chae, Y., & Davidson, T. (2025). Large Language Models for Text Classification: From Zero-Shot Learning to Instruction-Tuning. *Sociological Methods & Research*, 0(0). <https://doi.org/10.1177/00491241251325243>
- Chun, H. L., Lee, S. M., & Park, I. E. (2021). A systematic review of AI technology use in English education. *Multimedia-Assisted Language Learning*, 24(1), 87–103. <https://doi.org/10.15702/mall.2021.24.1.87>
- Firkins, A., Forey, G., & Sengupta, S. (2007). Teaching writing to low proficiency EFL students.

- ELT Journal*, 61(4), 341–352. <https://doi.org/10.1093/elt/ccm052>
- Fleckenstein, J., Liebenow, L. W., & Meyer, J. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1162454>
- Google for Educators. (2024b). *Generative*. Mountain View, CA: AI for educators. Google LLC. <https://skillshop.exceedlms.com/student/path/1176018>.
- Graham, S., Harris, K., & Hebert, M. (2011). *Informing writing: The benefits of formative assessment: A report from Carnegie Corporation of New York*. New York: Carnegie Corporation.
- Grimes, D. & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Learning and Assessment*, 8(6). <https://www.ejournals.bc.edu/index.php/jtla/article/view/1625/1469>
- Hasanah, U., & Safitri, W. (2024). Student perception toward ChatGPT in writing skill in English study program Universitas Muhammadiyah of Bengkulu. *Teaching and Learning English in the Language Education (Telle)*, 4(2). <https://doi.org/10.36085/telle.v4i2.6937>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Ho, N. N., Tra, T. C. T., Pham, T. Q. & Tran, P. T. (2026). Discrepancies between ChatGPT and Vietnamese EFL teachers in writing assessment. *International Journal of AI in Language Education*, 3(1), 1–16. <https://doi.org/10.54855/ijaile.26311>
- Jacobs, H. L., Zinkgraf, S. A., Wormouth, D.R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House. [ERIC ED217708](https://eric.ed.gov/?id=ED217708). <https://eric.ed.gov/?id=ED217708>
- Kurt, G., & Kurt, Y. (2024). Enhancing L2 writing skills: ChatGPT as an automated feedback tool. *Journal of Information Technology Education: Research*, 23, Article 24. <https://doi.org/10.28945/5370>
- Lại, T. K. L. & Trịnh, H. L. (2025). TÁC ĐỘNG CỦA CHATGPT ĐỐI VỚI NĂNG LỰC VIẾT HỌC THUẬT CỦA SINH VIÊN CHUYÊN NGÀNH TIẾNG ANH TẠI MỘT TRƯỜNG ĐẠI HỌC. *Tạp chí Nghiên cứu nước ngoài*, 41(3), 92-109. <https://doi.org/10.63023/2525-2445/jfs.ulis.5514>
- Lam, R. (2018). Feedback in writing portfolio assessment. In R. Lam (Ed.), *Portfolio assessment for the teaching and learning of writing*, (pp. 59–72). Springer. https://doi.org/10.1007/978-981-13-1174-1_5
- Ljubojević, D. (2024). Student feedback on the effectiveness of ChatGPT for essay writing. In *Proceedings of the 15th International Conference on e-Learning 2024* (pp. 116-126). CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3938/Paper_12.pdf
- 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>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(9). <https://doi.org/10.1186/s40561-024-00295-9>
- Meniado, J. C., Duong, T. T. H., Panyadilokpong, N., & Lertkomolwit, P. (2024). Using

- ChatGPT for second language writing: Experiences and perceptions of EFL learners in Thailand and Vietnam. *Computers and Education: Artificial Intelligence*, 7, 100313. <https://doi.org/10.1016/j.caeai.2024.100313>
- Nguyen, D. M. Q. (2026). The role of ChatGPT in providing written corrective feedback for EFL academic writing: A literature review. *International Journal of AI in Language Education*, 3(1), 35–47. <https://doi.org/10.54855/ijaile.26313>
- 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>
- Poláková, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11, 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
- Rahman, N., Zulkornain, H., & Hamzah, H. (2022). Exploring artificial intelligence using automated writing evaluation for writing skills. *Environment-Behaviour Proceedings Journal*, 7(SI9), 547–553. <https://doi.org/10.21834/ebpj.v7iSI9.4304>
- Setyowati, L., Sukmawan, S., & El-Sulukiyyah, A. A. (2020). Exploring the use of ESL composition profile for college writing in the Indonesian context. *International Journal of Language Education*, 4(2), 171–182. <https://doi.org/10.26858/ijole.v4i2.13662>
- Stillman, J. (2025, February 25). How to write the perfect AI prompt, according to OpenAI President Greg Brockman. *Inc.com*. <https://www.inc.com/jessica-stillman/how-to-write-perfect-ai-prompt-openai-greg-brockman.html>
- Shermis, M. & Wilson, J. (Eds). (2024). *Routledge international handbook of automated essay evaluation*. Routledge (in 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 of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Texas A&M University Writing Center. (2026). *Formative feedback*. Texas A&M University. <https://writingcenter.tamu.edu/faculty/resources/formative-feedback.html>
- Wang, M., & Guo, W. (2023). The potential impact of ChatGPT on education: Using history as a rearview mirror. *ECNU Review Education*. <https://doi.org/10.1177/20965311231189826>.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35, 24824–24837. https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf
- Wilson, J., Myers, M. C., & Potter, A. (2022). Investigating the promise of automated writing evaluation for supporting formative writing assessment at scale. *Assessment in Education: Principles, Policy & Practice*, 1–17. <https://doi.org/10.1080/0969594X.2022.2025762>
- Woo, D. J., Guo, K., & Susanto, H. (2023). *Cases of EFL secondary students' prompt engineering pathways to complete a writing task with ChatGPT*. arXiv.

<https://doi.org/10.48550/arXiv.2307.05493>

Zhang, R. & Chen, J. (2026). The effects of ChatGPT on EFL students' writing skills: A meta-analysis. *System*, 140, 104043. <https://doi.org/10.1016/j.system.2026.104043>

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>

Biodata

Ms. Nguyen Kim Phuc has been working as an English lecturer at the Faculty of Foreign Languages, Hanoi University of Science and Technology (HUST) since 2019. Her research in ELT primarily focuses on enhancing students' productive skills.

Ms. Vu Quynh Hoa has also been working as an English lecturer at the Faculty of Foreign Languages, Hanoi University of Science and Technology (HUST) since 2019. Her research interests are students' engagement, the application of technology, and EFL learners' writing skills.