


International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 1, pp. 94-110, January 2026
<https://doi.org/10.26803/ijlter.25.1.5>
Received Oct 10, 2025; Revised Dec 4, 2025; Accepted Dec 6, 2025

Reimagining Writing Feedback: The Effect of Copilot Artificial Intelligence (AI) on English Foreign Language (EFL) Students' Written Performance

Abdullah Al Maqbali* 
Sohar University, Oman

Salah Al-Hanshi 
Sultan Qaboos University, Oman

Ali Al Ghaithi 
Sohar University, Oman

Behnam Behforouz 
University of Technology and Applied Sciences, Shinas, Oman
Sohar University, Oman

Abstract. This study aimed to investigate the impact of using Copilot AI to provide corrective feedback on the written work of language learners and compare it with the corrective input from teachers. Using a quasi-experimental design, 70 Omani English language learners with an intermediate English proficiency level were assigned to a control group (which received general corrective feedback from the teacher) and an experimental group (which received detailed feedback from the Copilot AI and extra instructions from the teacher after Copilot AI), with 35 students in each group. Two comparison-contrast writing tasks were used as pretests and posttests and scored by two markers on Task Achievement, Organisation, Grammar and Lexicon. Although the initial findings revealed improvements in writing skills in both control and experimental groups, the detailed analysis of the posttest indicated that the experimental group performed significantly better than the control group in developing writing skills. In addition, the significant size effects of both tests revealed that both groups started with similar knowledge of writing skills in the pretest, but experimental group's performance showed larger effect sizes which reveal higher performance in written

*Corresponding author: Abdullah Al Maqbali; AMmaqkali@su.edu.om

tasks. The results of the study are insightful for teachers, learners and institutions.

Keywords: Artificial Intelligence feedback; corrective feedback; writing Skills

1. Introduction

Suleiman (2000) posits that writing is an essential element of language; therefore, every syllabus must incorporate the complex nature of writing into teaching methods, assessment strategies and language development. Gay (2002) found that writing is connected to improving students' critical thinking abilities and learning.

However, as writing is one of the most challenging aspects of learning a language, many researchers have been interested in exploring the nature of these issues (Banaruee, 2016; Richards & Renandya, 2002). Writing obligations require students to demonstrate sophisticated skills even when their proficiency is insufficient (Bitchener & Ferris, 2012; Khadawardi, 2020). Therefore, teachers need to invest a significant amount of time and effort in providing students with helpful comments on their writing (Zhang, 2017). On the other hand, feedback is essential for teaching how to write in a second language (Hyland & Hyland, 2019).

By engaging in feedback activities, learners enhance their psychological processes within their Zone of Proximal Development through cooperative learning with more skilled individuals (Fithriani et al., 2019). Research demonstrates considerable benefits of corrective feedback. Firstly, the corrective feedback process improves students' writing quality in terms of content, organisation, vocabulary, language use and mechanics (Farrak, 2012; Kahyalar & Yilmaz, 2016). Secondly, participating in corrective feedback activities helps students develop their critical thinking skills (Berg, 1999). Thirdly, corrective feedback encourages learner autonomy, enabling students to independently identify and revise their writing (Miao et al., 2006). Additionally, getting corrective comments on their work makes learners more aware of writing for an audience (Fithriani et al., 2019).

Nonetheless, several studies have demonstrated that both educators and students encounter challenges in providing and receiving effective corrective feedback. Students find it demoralising when their teachers criticise their work because the teachers notice many mistakes and struggle to suggest improvements. Moreover, students showed opposition, as students do not appreciate having to write repeatedly because they regard it as merely replicating their drafts (Lee, 2016).

Teachers play a key role in this process, primarily by correcting written mistakes and marking students' papers with the necessary corrections (Furieux et al., 2007). They also spend a lot of time grading numerous papers without realising the effectiveness of their methods (Lee, 2016). Moreover, the training of writing teachers regarding feedback is still insufficiently advanced (Ferris et al., 2011). In response to this challenging situation, AI can be viewed as a valuable resource,

since this innovative technology has the potential to alleviate teachers' substantial workload and deliver personalised feedback to improve active learning (Behforouz & Frumuselu, 2020; Kasneci et al., 2023).

Over the last few years, artificial intelligence has slowly become a part of language learning for writing. Research has shown that AI-powered tools provide rapid and individualised responses on grammatical and stylistic matters (Behforouz et al., 2025; Cahyono & Kurniawan, 2020; Tran, 2024). AI systems can offer rapid and scalable feedback for large courses, relieving teachers of the burden (Xiong et al., 2020). Miranty and Widiati (2021) stated that students who received AI input outperformed those who were provided with conventional feedback when comparing writing scores.

Additionally, AI allows for real-time feedback, providing students with instant, specific guidance and reducing administrative time to improve productivity and effectiveness of operations (Amin et al., 2025). Earlier research indicates that automatic writing feedback (AWF) can be used to assess, evaluate and grade as well as qualified human evaluators (Ramesh & Sanampudi, 2022). Also, it has been shown that dynamic feedback focused on students' strengths or areas for improvement can significantly impact writing and revision (Fleckenstein et al., 2023) and overall effectiveness (Huang & Wilson, 2021).

Although there have been many studies focusing on the implementation of Web 2.0 technologies in L2 studies, empirical research examining text prediction and AI writing tools is relatively rare. In addition, the point which is not clearly shown in the existing literature is to understand how the AI-generated feedback works beside the teacher's feedback to improve the writing skills of the learners, particularly in English language learning context. As a result, it is necessary to explore the possible applications of AWF as an additional input in language classrooms aimed at improving students' writing skills (Frankenberg-Garcia, 2020; Mustopa et al., 2024). Furthermore, earlier works in language learning and technology have examined the effectiveness of different applications and practices on language acquisition (Zhang & Zou, 2020). However, research on the use of AI-powered tools for higher education language acquisition is scant (Baskara & Mukarto, 2023).

Therefore, the goal of this study is to address the growing need for effective corrective feedback in EFL writing classrooms, where teachers often have a lot of work to do but insufficient time to provide individualised feedback. Additionally, it is an attempt to determine if using Copilot AI for full feedback may help Omani intermediate EFL students write better after receiving comments from an instructor alone. The results are significant because they show that AI technology can be used in language instruction to make learning more efficient, help teachers give consistent feedback, and encourage students to enhance their writing skills and independence in the classroom.

The study aimed at finding an answer to the following research question:

1. Does Copilot AI have any significant effect on the improvement of Omani intermediate EFL learners' writing skills?

2. Literature Review

2.1 Corrective Feedback by Teacher

Sakir et al. (2025) conducted a study examining teacher corrective feedback methods and student preferences with 15 learners in Indonesia. Teachers employed three of Rod Ellis's six feedback types: direct written corrective feedback (17 times), indirect written corrective feedback (15 times), and metalinguistic written corrective feedback (8 times). The replies from the students indicated that they would prefer to receive direct feedback over other ones. The strengths of the study included a realistic classroom setting and the inclusion of student perspectives; however, the drawbacks were a limited sample size and a qualitative methodology that lacks statistical analysis.

In another study, Wondim et al. (2024) conducted a quasi-experimental study examining the impact of written corrective feedback (WCF) on the writing abilities of 135 first-year Ethiopian college students. They were categorised into three groups: direct WCF with metalinguistic explanation, indirect WCF and a control group that received no feedback. For four weeks, the teacher gave constant, unfocused comments on all correctable mistakes in three writing sessions.

In the posttests, the direct WCF group did far better than the indirect and control groups. The study's benefits included ongoing feedback through several posttests and ecological validity in genuine classroom settings. However, the drawbacks involved a limited duration (4 weeks), a lack of long-term follow-up, and confinement to a specific writing genre. The findings of the study align with the Activity Theory as various types of corrective feedback act as intermediating instruments that direct the engagements of the learners within the written tasks and effect their improvements.

Unlike the previously described approaches, Al Ghaithi and Behforouz (2023) conducted an experimental study investigating the effect of corrective feedback on the writing abilities of 94 Omani EFL pre-intermediate learners in an online learning environment. The treatment group received weekly individual and group corrective feedback via Microsoft Teams for 12 weeks, while the control group received just general comments on form and structure of their written tasks.

The study found statistically significant differences favouring the experimental group. The study's strengths include an adequate sample size and random assignment; nevertheless, its weaknesses consist of a single-institution context, and an exclusively online environment that hinders generalisability. The findings could be related to the Noticing Hypothesis as the individual and group feedback to the experimental group could improve the abilities of the students in their writing skills which led to the transition of input to linguistic output.

2.2 The Impact of AI-Powered Tools on EFL Students' Writing Performance

Al Ghaithi and Behforouz (2025) investigated the impact of an AI-driven instrument on the writing achievement and perceptions of Omani EFL learners. Sixty pre-intermediate students were divided into three groups: control (general feedback with abbreviations only), experiment A (detailed one-on-one feedback from the researcher), and experiment B (feedback exclusively using EditGPT). All groups got the same lessons in class.

The experiment A group performed significantly better than both groups, while experiment B group outperformed the control, demonstrating that EditGPT is a valuable supplementary tool for enhancing text editing. Strengths include a comprehensive three-group comparison; weaknesses pertain to a restricted sample size and a brief four-week duration. The findings align with the Emergent AI-Assisted Discourse Theory, as the Human-AI interactions through EditGPT provide a supportive environment for the revisions of the text and improve the writing abilities through AI-mediated feedback.

Pratama et al. (2025) investigated the effect of artificial intelligence-powered instruments on the argumentative essay-writing skills of 60 Indonesian EFL students, employing a quasi-experimental design. Participants were divided into two groups: a treatment group that received instructional assistance from QuillBot and a control group that received traditional training. The study revealed that the performance in writing skills, coherence and grammatical accuracy was significantly improved in the treatment group compared to the control group. The strengths of the research include its thorough statistical analysis, while its limitations are a small sample size, reliance on a single educational institution, and the lack of qualitative data on student perceptions regarding the research. The results could be connected with the activity theory as the AI tool, QuillBot, acts as an intermediate tool affecting the performance of the students via AI-supported instruments.

Apriani et al. (2024) investigated the effect of an AI chatbot on writing skills, self-efficacy and self-regulation among 40 Indonesian EFL students using a mixed-methods design. Participants were assigned to either a control group that received traditional training or a treatment group that received chatbot-assisted writing. The results revealed that the treatment group made significant gains in writing ability, self-efficacy and internal control of behaviours, compared to members of the control group. Interviews indicated that such tools led to increased confidence, motivation and engagement among students, along with concerns about overreliance and the potential for plagiarism. The limitations of this study include its dual approach, which combines both qualitative and quantitative methods. It only studies at one university and has a limited sample. The findings align with the activity theory. The chatbot creates continuous interactive communications which forms the writing behaviour of the learners, confidence and self-regulation skills via AI-mediated engagement.

Oktarin et al. (2024) expanded the field of AI chatbots research by investigating how ChatGPT impacted the writing skills, motivation and feedback knowledge

of 50 Indonesian EFL learners. Over 16 weeks, students who had completed an Economics English course were trained in two groups, using either ChatGPT or a traditional teaching method. Writing results, feedback literacy and engagement were all significantly better for the treatment group than for the control group. Participants used strategies such as interaction practice, personal learning, peer assessment and gamified learning to develop these features. The limitations of this study include the small sample size (with only 50 women) and the focus on a single faculty, so generalizability is limited. The findings of this study are linked to the activity theory. It could be stated that the consistent interaction of the learners with ChatGPT builds an active learning environment to support the writing development, feedback and the learning engagement.

3. Method

This study was based on a quasi-experimental approach. This section outlines the study population, the data-collection instruments, the steps students followed when using Copilot AI for feedback, and the statistical analysis conducted with SPSS.

3.1 Participants

The sample population of this study consisted of 70 Omani English language learners from the Foundation Department at one of Oman's higher education institutions. These students, based on the institute's placement test, were placed in the intermediate level of English proficiency. All of them were native Arabic speakers whose ages ranged from 19 to 21.

Foundation Studies are mandatory steps for students before advancing to higher education specialisations based on the Oman National Curriculum. This period, which varies depending on individual universities, typically takes 1 to 2 year(s) to complete, starting from elementary levels to upper-intermediate ones. Students are instructed in courses such as English, Information Technology, Mathematics and occasionally Science.

Following a quasi-experimental study design, the participants in this study were assigned to two groups, each comprising 35 students with a combination of men and women. The control group followed the regular in-class teaching and learning on writing skills development, receiving general corrective feedback from their teachers. In contrast, the experimental group received similar in-class training on writing development and received detailed feedback from the Copilot AI. It is worth mentioning that due to the nature of the study, there was no control over age and gender. Additionally, monitoring of the students' access to other AI-powered tools was not considered.

3.2 Instruments

This section summarises the instruments used to collect the required data from both groups.

3.2.1 Writing tests package

To collect the required data from both groups and measure the impact of the treatment, a pretest and a posttest of writing, focusing on comparison and

contrast, were designed and conducted. Table 1 shows the written tasks that were selected for this purpose.

To score the tests, two researchers marked the students and assessed the writing tasks against four criteria: Task Achievement (5 marks), Organisation (5 marks), Grammar (5 marks), and Lexicon (5 marks), totalling 20 marks. The discrepancy between the first and second markers should not exceed 3 marks. In cases of higher discrepancies, the third marker may score the tasks. However, in this study, the scores fell within the discrepancy range, and therefore, there was no need for a third marker.

Table 1: The writing questions of the pretests and posttests

Instructions
<p>You have 40 minutes to write a comparison-contrast essay on the topic below. Your essay should include an introduction, two body paragraphs and a conclusion. Make sure to include clear points of similarity and difference. Use linking words such as 'similarly', 'on the other hand', 'however', 'both', and 'in contrast' to organise your ideas.</p> <p>Pretest Compare and contrast studying online and studying in a traditional classroom. Consider factors such as interaction, flexibility and learning experience.</p> <p>Posttest Compare and contrast living in a big city and living in a small town. Discuss aspects such as lifestyle, opportunities and environment.</p>

3.3 Procedure

The intervention took place during four weeks in the fall semester of 2025–2026. Prior to the study, all students gave their informed consent, knowing that their participation was voluntary. The week before the intervention, all the participants took a pretest to ensure the homogeneity of their writing skills. After that, the students in the experimental group had orientation sessions where they were guided on using Copilot AI through WhatsApp to help them write better. A customised Microsoft Teams platform made it easy for students to share materials, have asynchronous discussions and solve technical problems.

To ensure the lessons were the same for both groups, the related teachers had to take a two-day professional development course. The training covered the technical aspects of Copilot AI through WhatsApp, how to incorporate courses into the classroom, and how to ensure lessons are taught regularly. The intervention focused on one rhetorical mode: a comparison-contrast essay. The teaching sequence ensured that all groups were the same, but the methods of feedback were different.

In the experimental group, participants received consistent classroom instruction on writing conventions and rhetorical strategies. The students used Copilot AI through WhatsApp during in-class writing tasks to get quick feedback, which was then followed up by teacher clarification and direct responses. The first week was about the introduction and writing the first body paragraph. The instructor spent 30 minutes on each session explaining the structure of introductions to

comparison-contrast essays, including the hook, background and thesis statements. In the following 30-minute writing exercise, students used WhatsApp to get instant feedback from Copilot AI and then used what they had learnt in the session to improve their writing draft. After this, the teacher set aside 30 minutes to provide additional help on any confusing term raised from using Copilot AI. This strategy continued for the entire treatment period. Each week, the teacher introduced one part of the comparison-contrast essay until the full draft was covered, while students were receiving feedback from Copilot AI.

In the control group, students participated in structured classroom sessions where teachers offered direct instruction on essay form, rhetorical strategies and linguistic characteristics pertinent to each writing mode. In the first session, the teacher taught for 30 minutes about how to write effective introductory paragraphs for comparison-contrast essays, covering the hook, background and thesis statement. Then, for 30 minutes in class, the students wrote introduction paragraphs for different topics.

Thereafter, the teacher gave general feedback to the students for 30 minutes. During that time, examples of common mistakes from student papers, shown anonymously, were discussed to encourage group learning. This cycle also continued for the control group in the same way until the end of the treatment period and the completion of the comparison and contrast writing draft. At the end of the treatment period, learners from both groups participated in a writing posttest to compare their performance with the pretest.

3.4 Research Results

This section focuses on the statistical analysis of quantitative data using SPSS 27.0. Before conducting any test for comparison purposes, it was necessary to determine the data distribution condition among the participants in both pretests and posttests to choose either parametric or nonparametric tests for further analysis. Table 2 shows the results of the Shapiro-Wilk Normality Test.

Table 2: The results of data normality in pretests and posttests of both groups

	Groups	Shapiro-Wilk		
		Statistic	df	Sig.
Pretest	Control	.950	35	.113
	Experiment	.951	35	.118
Posttest	Control	.965	35	.320
	Experiment	.945	35	.079

Table 2 shows that in the control group ($p = .113$) and the experimental group ($p = .118$) the data distributions were approximately normal. In the posttest, the control group ($p = .320$), and the experimental group ($p = .079$) showed p -values greater than 0.05; therefore, the data were roughly normally distributed, and parametric tests could be used for further comparison of the groups together. To analyse the performance of students within their groups, paired-samples t -tests were used, and Table 3 shows the results of these paired comparisons.

Table 3: The performance of learners within their groups in pretest and posttest

		Paired Differences								
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)	
					Lower	Upper				
control	pretest - posttest	-3.714	2.269	.383	-4.493	-2.934	-9.682	34	.000	
experiment	pretest - posttest	-7.000	2.765	.467	-7.949	-6.050	-14.976	34	.000	

Table 3 shows that the control group with a mean difference of -3.71 ($SD = 2.27$) $t(34) = -9.68$, $p < .001$ showed significant performance over the time. The experimental group with a mean difference of -7.00 ($SD = 2.77$), $t(34) = -14.98$, $p < .001$ also showed significant performance in the posttest. The 95% confidence intervals for the mean differences (Control: $[-4.49, -2.93]$; Experimental: $[-7.95, -6.05]$) showed that both improvements were statistically significant. As both groups showed progress within their own groups, a one-way ANOVA was therefore undertaken to compare the performance of learners from both groups in the pretest and posttest. Table 4 shows the results of such analysis.

Table 4: The results of comparing the pretest and post test of both groups

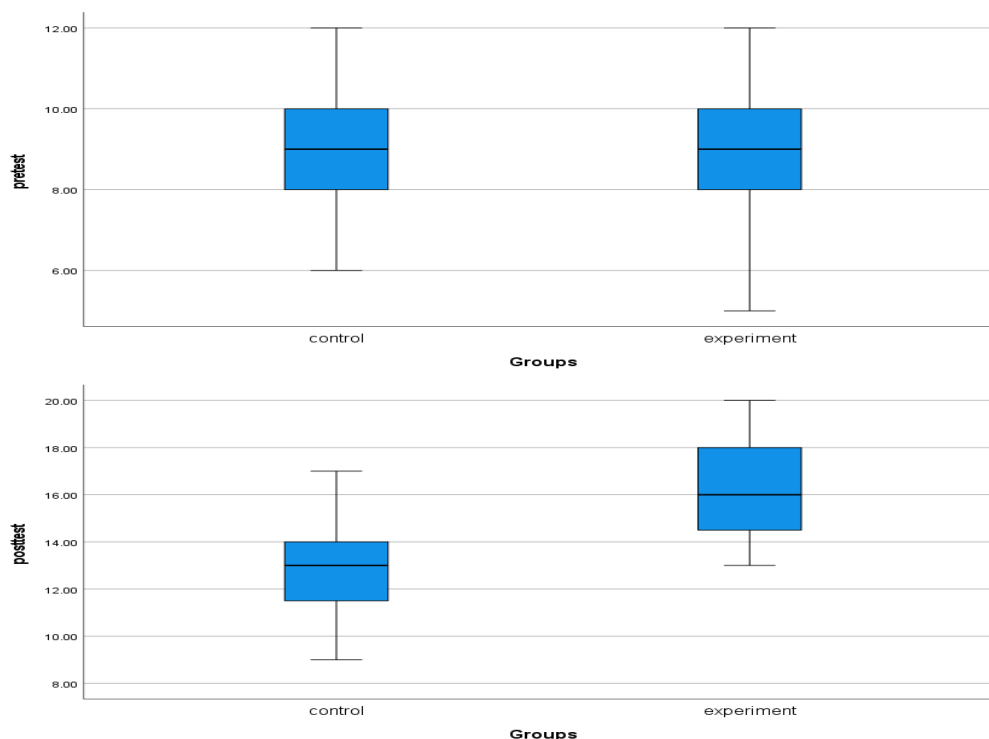
		Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	.229	1	.229	.108	.743
	Within Groups	143.714	68	2.113		
	Total	143.943	69			
Posttest	Between Groups	205.714	1	205.714	50.000	.000
	Within Groups	279.771	68	4.114		
	Total	485.486	69			

Table 4 shows that in the pretest, there was no statistically significant difference between the two groups, $F(1, 68) = 0.108$, $p = .743$. This means that both groups had the same level of writing skills at the start. The posttest results showed a statistically significant difference between the groups, $F(1, 68) = 50.00$, $p < .001$, with the experimental group achieving significantly higher writing scores than the control group. The statistics show that, although the groups started at similar levels, the experimental group, which received feedback based on Copilot AI, demonstrated a significant improvement in their writing skills by the end of the treatment. To determine the magnitude of this discrepancy, the effect sizes of the scores were measured in the posttest, and the results are presented in Table 5 below.

Table 5: The results of effect size measures in the posttest between the groups

Posttest		Point Estimate	95% Confidence Interval	
			Lower	Upper
	Eta-squared	.424	.245	.555
	Epsilon-squared	.415	.233	.548
	Omega-squared Fixed-effect	.412	.231	.545
	Omega-squared Random-effect	.412	.231	.545

Table 5 shows that the eta-squared value was 0.424, with a 95% confidence interval of 0.245 to 0.555. This means that the treatment conditions explained around 42.4% of the difference in writing competence. In the same way, the epsilon-squared value (.415, 95% CI [.233, .548]) and both the fixed- and random-effects omega-squared values (.412, 95% CI [.231, .545]) showed that there was a big impact size. Practically, the revealed significant size shows that the benefits of using Copilot AI to receive feedback on the written tasks could be observed in real classes too. From the teachers' perspectives, the use of AI-generated feedback could improve the writing skills of the learners, reduce the efforts in providing individualised feedback, and offer consistent learning support across the tasks. To present the data from pretests and posttests visually, Figure 1 compares the score distributions of both groups in two sets of tests.

**Figure 1: The distribution of scores of both groups in pretests and posttests**

4. Discussion

The primary objective of this study was to measure the impact of receiving feedback from the Copilot AI on the writing skills of English learners. For this purpose, 70 Omani intermediate EFL students were assigned to a control group

and an experimental group, with 35 students in each group. While the control group received regular in-class training on writing development and feedback from the class teacher, the experimental group received similar in-class training but received their feedback from the Copilot AI. After conducting pretests and posttests of writing for a period of one month, the initial findings showed an increase in the writing skills of students in both groups; however, further thorough analysis of the posttest revealed that the students of the experimental group which received detailed corrective feedback from Copilot AI, significantly outperformed the control group. This could be strongly associated with the use of AI-powered tools within the language learning process to receive detailed feedback on their written products.

The significant performance of the experimental group could be related to the advantages of AI-mediated feedback. Copilot AI's features offered fast, accurate and comprehensive analysis of errors, from correction of grammar to style recommendations, which cannot be done regularly by the teachers. This feedback increases the possibility of detecting linguistic errors and facilitates the processes of revision (such as error correction and rewriting). In contrast, the control group, which relied solely on their teachers' feedback, would have engaged in fewer rounds of corrective revision and obtained less immediate corrective feedback. The feedback environment in the experimental group was more robust with quicker feedback, greater coverage of corrective recasts and opportunities for self-correction and metalinguistic reflection by students.

AI-powered tools feedback provided by the large language models provides some significant implications for the traditional techniques that reshape the relationship between teachers and students and contributes to the learner's autonomy. Although teachers traditionally provided feedback, evaluation and instructions for the students, AI-powered tools are working as facilitators in the learning process to foster immediate and customised feedback. Simultaneously, this digital shift aligns with the learner-centred approaches by encouraging independent revision, language experiments and self-regulated learning engagement. There are some theories associated closely with these findings. Activity theory functions as a foundational theoretical framework that synthesises these findings, highlighting the mediating role of objects (such as Copilot AI) in learning processes within sociocultural systems.

Rahimi et al. (2024) articulated this viewpoint in their study of automated WCF, illustrating that AI-generated feedback facilitates the internalisation of revision and, and operates as a co-mediator in the learner's interaction with texts. From this perspective, AI-generated feedback facilitates students' participation in more sophisticated review procedures. Jacob et al. (2023) asserts that collaboration between individuals and AI-powered tools in writing may produce innovative discursive patterns that increase flexibility for learners. The study demonstrated that AI-enhanced technology did not supplant learner voice; instead, it enabled novel rhetorical techniques and transformations. The results are consistent with the Noticing Hypothesis (Schmidt, 1990), which asserts that learners must identify inconsistencies between their interlanguage and target norms to internalise new

structures. AI feedback offered students clear instructions on how to fix their mistakes, which helped them become more aware of language and learn more deeply. The two frameworks together give a theoretical reason for why the writing results were better in experimental conditions. The input from Copilot AI made it easier for learners to think and write in a way that helped them become more aware and scaffolded their work so they could keep making progress. The present findings substantiate the concept that AI feedback serves not merely as a mechanism for error correction but also as an active participant in the educational process, augmenting students' writing proficiency beyond the constraints of conventional teacher feedback.

The study's results align with Duong et al. (2025), who investigated the involvement of EFL students with the Writing Assistant Bot (WAB) to enhance writing proficiency. The research demonstrated that the chatbot significantly improved students' writing skills in all areas, including content, organisation, vocabulary, language use and mechanics. In line with these findings, Mahapatra (2024) investigated the impact of AI on the writing capabilities of undergraduate students, reporting a significant favourable effect on their academic writing proficiency.

Polakova and Ivenz (2024) investigated the effectiveness of AI feedback in improving students' writing skills, demonstrating improvements in several areas, including conciseness, grammatical accuracy, inclusion of pertinent information and the use of passive voice. Al Mahmud (2023) investigated the extent to which AI-driven technologies facilitate writing among Saudi students and found that the experimental group that used the AI-powered tool made more progress than the control group. Additionally, Song and Song (2023) evaluated the impact of AI-assisted language learning on the writing proficiency of Chinese students, demonstrating significant improvements in writing skills and motivation among participants receiving AI-assisted instruction compared to the control group.

5. Conclusion

This study aimed to investigate the impact of using Copilot AI to provide corrective feedback for English language learners in Oman. The study's results showed progress from pretest to posttest, supporting the benefits of receiving corrective feedback from both providers. However, the comparison between the two groups showed that the experimental group, which received Copilot AI feedback, performed significantly better than the control group.

The study's recommendations are helpful for teachers, students and schools. By using Copilot AI feedback, teachers may be able to lighten their load while still providing high-quality and accurate extra help. Instead of fixing the same mistakes many times, teachers may spend more time teaching advanced writing skills, such as organizing arguments and ideas clearly or writing in an academic manner. Additionally, the AI system helps teachers make better classes and formative tests. The method gives students quick, personalized and ongoing feedback, which encourages them to learn on their own, practice more and repair their own mistakes. They use Copilot AI to help students make regular and

consistent changes to their language, which helps them write more accurately, expressively and fluently. AI feedback keeps students motivated and engaged by giving them constant feedback and cutting down on the time it takes for teachers to provide feedback. The use of AI-enhanced feedback systems encourages new ways of teaching and using technology in the classroom, which is in line with current trends. It makes learning more effective and makes it easier to judge how well education is working. The school is open to using AI in language education, which improves its academic reputation and draws in students and teachers who are interested in technology. In total, the educators and educational institutions could integrate the AI-powered tools as learning facilitators to enhance the human instructions rather than replacing them.

In the classroom settings, AI-powered tools could be assisting with the first drafting, task revising, and helping the learners to compare the AI suggestions with the teachers points to foster critical judgements and self-regulation skills. Additionally, the teachers should receive training sessions on command design, AI literacy and ethical usage to align the objectives with the learning outcomes and academic integrity. Institutions could provide policies on AI usage, provide students with educational platforms, and embed the AI-powered tools with their writing centres to create a balance environment where the AI-powered tools act as facilitators beside the teacher's guidance and accountability.

The present study possesses specific methodological and contextual deficiencies that necessitate recognition. The four-week length of the intervention may not have been long enough to cause significant changes, which makes it harder to figure out how AI-based feedback affects writing development over time. The short period of the treatment could not provide information regarding the long-term effects of AI-powered tools on developmental progress but focusing on the immediate performance gains. To measure such developments, continuous practice, exposure to the AI-powered tools, and the opportunities to use the new writing strategies and techniques are in need in longer period of the treatment.

Secondly, the participants had an intermediate level of English skill, making it difficult to apply the results to higher levels of English proficiency. The study concentrated solely on the writing-specific AI program, Copilot AI, and did not assess the potential impacts of similar technology on other linguistic skills, including speaking, listening or reading. The technique adopted was solely quantitative, excluding qualitative data, such as interviews or reflective diaries, which could provide a more comprehensive insight into learners' thoughts and experiences. The lack of control over demographics such as age and gender may have affected the results, as these factors can alter the receptiveness to technology-enhanced input.

Future research can mitigate these limitations by prolonging the treatment duration to investigate the long-term effects of AI-based feedback beyond a single academic semester or school year. Re-tasking may also involve including students with varying levels of English proficiency to assess whether AI feedback has a disparate effect on language skills. Expanding the scope of this study to include

more AI applications could produce similar results concerning the efficacy and reliability of remedial feedback. Furthermore, using qualitative methodologies, including semi-structured interviews, learner diaries or focus groups, could provide a deeper understanding of students' engagement, perspectives and emotions around AI-assisted learning. Demographic characteristics, such as age and gender, have a significant effect on educational achievements; therefore, future studies could focus on these factors for the purposes of comparisons. Ultimately, these and similar experimental setups can be used to improve linguistic skills, truly understanding how AI technologies might enhance overall language competency.

6. Acknowledgements

The authors wish to acknowledge the use of [Grammarly AI] in the proofreading of the first draft of this manuscript. This tool was used to help improve the language and grammar in the paper.

7. References

- Al Ghaithi, A., & Behforouz, B. (2025). Effects of interaction with AI-Assisted writing evaluation on EFL students' writing performance. *Knowledge Management & E-Learning*, 17(2), 206–224. <https://doi.org/10.34105/j.kmel.2025.17.009>
- Al Ghaithi, A., & Behforouz, B. (2023). The effect of corrective feedback via a computerized course on Omani EFL learners writing performance. *Turkish Online Journal of Distance Education*, 24(1), 74–87. <https://doi.org/10.17718/tojde.1025572>
- Al Mahmud, F. (2023). Investigating EFL students' writing skills through artificial intelligence: Wordtune application as a tool. *Journal of Language Teaching and Research*, 14(5), 1395–1404. <https://doi.org/10.17507/jltr.1405.28>
- Amin, M. R. M., Ismail, I., & Sivakumaran, V. M. (2025). Revolutionizing education with Artificial Intelligence (AI)? Challenges, and implications for open and distance learning (ODL). *Social Sciences & Humanities Open*, 11, 01308. <https://doi.org/10.1016/j.ssaho.2025.101308>
- Apriani, E., Cardoso, L., Obaid, A. J., Muthmainnah, M., Wijayanti, E., Esmianti, F., & Supardan, D. (2024). Impact of AI-powered chatbots on EFL students' writing skills, self-efficacy, and self-regulation: A mixed-methods study. *Global Educational Research Review*, 1(2), 57–72. <https://doi.org/10.71380/GERR-08-2024-8>
- Banaruee, H. (2016). *Recast in writing*. Sana Gostar Publications.
- Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 7(2), 343–358. <http://dx.doi.org/10.21093/ijeltal.v7i2.1387>
- Behforouz, B., & Frumuselu, A. D. (2020). The impact of text messaging as an instructional tool to enhance learner autonomy and perception. *International Journal of Learning, Teaching and Educational Research*, 19(11), 184–202. <https://doi.org/10.26803/ijlter.19.11.11>
- Behforouz, B., Al Ghaithi, A., & Al Balushi, Z. (2025). Revolutionizing vocabulary acquisition through AI chatbots. *International Journal of Learning, Teaching and Educational Research*, 24(7), 1–22. <https://doi.org/10.26803/ijlter.24.7.1>
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215–241. [https://doi.org/10.1016/S1060-3743\(99\)80115-5](https://doi.org/10.1016/S1060-3743(99)80115-5)
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge. <https://doi.org/10.4324/9780203832400>

- Cahyono, B. Y., & Kurniawan, D. A. (2020). Evaluating the effectiveness of an AI-powered assessment tool for writing. *Journal of English Language Teaching and Linguistics*, 5(2), 217–232.
- Duong, T.-N.-A., & Chen, H.-L. (2025). An AI chatbot for EFL writing: Students' usage tendencies, writing performance, and perceptions. *Journal of Educational Computing Research*, 63(2), 406–430. <https://doi.org/10.1177/07356331241312363>
- Farrah, M. (2012). The impact of peer feedback on improving the writing skills among Hebron university students. *An - Najah University Journal of Research - Humanities*, 26(1), 179–210.
- Ferris, D., Brown, J., Liu, H. (Sean), & Stine, M. E. A. (2011). Responding to L2 students in college writing classes: Teacher perspectives. *TESOL Quarterly*, 45(2), 207–234. <https://doi.org/10.5054/tq.2011.247706>
- Fithriani, R., Rafida, T., & Siahaan, A. (2019). Integrating online blogging into EFL writing instruction: Exploring students' perceptions. *Proceedings of the UNNES International Conference on English Language Teaching, Literature, and Translation (ELTLT 2018)*, 188, 87–90. <https://doi.org/10.2991/eltlt-18.2019.17>
- 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, 1162454. <https://doi.org/10.3389/frai.2023.1162454>
- Frankenberg-Garcia, A. (2020). Combining user needs, lexicographic data and digital writing environments. *Language Teaching*, 53(1), 29–43. <https://doi.org/10.1017/S0261444818000277>
- Furieux, C., Paran, A., & Fairfax, B. (2007). Teacher stance as reflected in feedback on student writing: An empirical study of secondary school teachers in five countries. *IRAL-International Review of Applied Linguistics in Language Teaching*, 45, 69–94. <http://dx.doi.org/10.1515/IRAL.2007.003>
- Gay, G. (2002). *Developing dimensions of written language*. Adaptive Technology Resource Centre.
- Huang, Y., & Wilson, J. (2021). Using automated feedback to develop writing proficiency. *Computers and Composition*, 62, 102675. <https://doi.org/10.1016/j.compcom.2021.102675>
- Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In *feedback in second language writing: Contexts and issues* (pp. 1–22). Cambridge University Press. <https://doi.org/10.1017/9781108635547.003>
- Jacob, S., Tate, T., & Warschauer, M. (2023). *Emergent AI-Assisted discourse: Case study of a second language writer authoring with chatGPT*. arXiv. <https://arxiv.org/abs/2310.10903>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kahyalar, E., & Yılmaz, F. (2016). Teachers' corrective feedback in writing classes: The impact of collaborating with a peer during the editing process on students' uptake and retention. *The Reading Matrix: An International Online Journal*, 16(1), 148–160.
- Khadawardi, H. A. (2020). The effect of implicit corrective feedback on English writing of international second language learners. *English Language Teaching*, 14(1), 123–139. <https://doi.org/10.5539/elt.v14n1p123>
- Lee, I. (2016). Teacher education on feedback in EFL writing: Issues, challenges, and future directions. *TESOL Quarterly*, 50(2), 518–527. <https://doi.org/10.1002/tesq.303>

- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Miranty, D., & Widiati, U. (2021). An automated writing evaluation (AWE) in higher education. *Pegem Journal of Education and Instruction*, 11(4), 126–137. <https://doi.org/10.47750/pegegog.11.04.12>
- Mustopa, Nasikhin, Chamami, R., Nihayah, H., Habibullah, M. R., & Manshur, A. (2024). Challenges in artificial intelligence development in higher education in China, India, and Indonesia: International students' perspectives. *International Journal of Learning, Teaching and Educational Research*, 23(2), 354–373. <https://doi.org/10.26803/ijlter.23.2.17>
- Oktarina, I. B., Saputri, M. E. E., Magdalena, B., Hastomo, T., & Maximilian, A. (2024). Leveraging ChatGPT to enhance students' writing skills, engagement, and feedback literacy. *Edelweiss Applied Science and Technology*, 8(4), 2306–2319.
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
- Pratama, Y., Fridolini, F., & Pitaloka, R. M. (2025). The impact of AI Quillbot in improving student writing ability to write argumentative essays. *English Review: Journal of English Education*, 13(1), 215–222. <https://doi.org/10.25134/erjee.v13i1.9746>
- Rahimi, M., Fathi, J., & Zou, D. (2025). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies*, 30, 2691–2735. <https://doi.org/10.1007/s10639-024-12896-5>
- Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring system: A systematic literature review. *Artificial Intelligence Review*, 55, 2495–2527. <https://doi.org/10.1007/s10462-021-10068-2>
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge university press. <https://doi.org/10.1017/CBO9780511667190>
- Sakir, R., Sunra, L., & Korompot, C. A. (2025). Teacher corrective feedback in students' English writing at SMAN 3 Selayar. *ARRUS Journal of Social Sciences and Humanities*, 5(1), 775–786. <https://doi.org/10.35877/soshum3546>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Suleiman, M. F. (2000). *The process and product of writing: Implications for elementary school teachers*. 2–11. <https://eric.ed.gov/?id=ED442299>
- Tran, T. T. H. (2023). AI tools in teaching and learning English academic writing skills. *Proceedings of the AsiaCALL International Conference*, 4, 170–187. <http://asiacall.info/proceedings/index.php/articles/article/view/86>
- Wondim, B. M., Bishaw, K. S., & Zeleke, Y. T. (2024). Effectiveness of teachers' direct and indirect written corrective feedback provision strategies on enhancing students' writing achievement: Ethiopian university entrants in focus. *Heliyon*, 10(2). <https://doi.org/10.1016/j.heliyon.2024.e24279>
- Xiong, Z., Ma, Y., Wu, Y., & Liu, J. (2020). Automatic essay grading using machine learning. *International Journal of Emerging Technologies in Learning (iJET)*, 15 (17), 63–78.

- Zhang, X. (2017). Reading-writing integrated tasks, comprehensive corrective feedback, and EFL writing development. *Language Teaching Research*, 21(2), 217-240. <https://doi.org/10.1177/1362168815623291>
- Zhang, R., & Zou, D. (2022). Self-regulated second language learning: A review of types and benefits of strategies, modes of teacher support, and pedagogical implications. *Computer Assisted Language Learning*, 37(4), 720-765. <https://doi.org/10.1080/09588221.2022.2055081>