



Castledown

 OPEN ACCESS

# Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>


*Technology in Language Teaching & Learning*, 7(3), 102626 (2025)

<https://doi.org/10.29140/tltl.v7n3.102626>

## Comparing Teacher E-Feedback, AI Feedback, and Hybrid Feedback in Enhancing EFL Writing Skills



AFSHIN SOORI<sup>a</sup> 

LALEH KHOJASTEH<sup>b</sup> (Corresponding Author) 

FAREEHA JAVED<sup>c</sup>

<sup>a</sup>*Department of English Language and Literature,  
Islamic Azad University, Larestan, Iran  
afshin\_soori@yahoo.com*

<sup>b</sup>*Department of English Language, School of  
Paramedical Sciences, Shiraz University of  
Medical Sciences, Shiraz, Iran  
khojastehlaleh@yahoo.com*

<sup>c</sup>*Department of English Language & Literature,  
Government Graduate College for Women, Gulberg, Lahore Pakistan  
fjaved71@gmail.com*

### Abstract

Effective feedback plays a critical role in enhancing the writing skills of English as a Foreign Language (EFL) learners. This study examines the comparative effectiveness of three feedback approaches—Teacher e-feedback, AI-based feedback, and a hybrid model—in enhancing the writing performance of Iranian intermediate-level EFL learners. A randomized controlled trial was conducted with 88 intermediate-level EFL learners, who were randomly assigned to one of three feedback groups: (1) teacher e-feedback, (2) AI-generated feedback using tools such as ChatGPT and Grammarly, and (3) a hybrid approach that combined both feedback types. Writing proficiency was assessed using IELTS writing tasks and the Oxford Placement Test before and after the intervention. Significant differences were observed between the groups, with the hybrid feedback group showing the most substantial improvements, particularly in task achievement, coherence, and grammatical accuracy. AI feedback was most effective in enhancing lexical resources. Qualitative reflections supported the quantitative

**Copyright:** © 2025 Afshin Soori, Laleh Khojasteh & Fareeha Javed. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.  
**Data Availability Statement:** All relevant data are within this paper.

findings, with participants in the hybrid group reporting increased confidence, reduced anxiety, and appreciation for the balanced, dual-source feedback. These results highlight the pedagogical potential of integrating human and AI feedback to enhance EFL writing instruction.

Keywords: EFL writing; teacher e-feedback; AI feedback; hybrid feedback; writing performance

---

## Introduction

In the context of English language learning, particularly for English as a Foreign Language (EFL) and English as a Second Language (ESL) learners, writing is often considered one of the most challenging skills to master. Effective writing requires a deep understanding of grammar, the ability to structure coherent and cohesive arguments, and the capability to respond appropriately to a variety of tasks. These complexities make writing instruction a critical area of focus in language education. However, students frequently struggle with mastering these elements, facing challenges such as grammatical inaccuracies, disorganized ideas, and difficulty in fully achieving the objectives of a writing task (Febriani, 2022). This anxiety can hinder students' willingness to engage in writing tasks, ultimately affecting their learning outcomes (Tsao et al., 2021).

Considering these challenges, feedback plays a pivotal role in writing instruction. Effective feedback provides learners with the guidance needed to identify and correct errors, refine their ideas, and develop clearer, more structured arguments (Wulandari, 2022). It serves as a crucial tool in helping students bridge the gap between their current performance and their potential capabilities (Yamalee & Tangkiengsirisin, 2019). The quality and delivery method of feedback are essential factors that influence its effectiveness. Traditional teacher feedback, renowned for its contextual richness and personalization, often demands considerable time and effort from educators and may not always be delivered promptly (Erkan, 2022). On the other hand, recent technological advancements have introduced AI-based feedback systems, which can provide immediate and detailed responses to students' writing. These systems excel at addressing surface-level issues, such as grammar and syntax, by offering instant corrections that students can quickly act upon (Wilson & Roscoe, 2020).

The rise of AI in educational settings has sparked an ongoing debate among educators and researchers regarding the most effective approach to feedback. Some advocate for maintaining the traditional, human-driven feedback that leverages the teacher's expertise and ability to offer nuanced, empathetic responses tailored to each student's needs (Escalante et al., 2023). Others argue for embracing AI-based feedback due to its efficiency and the immediacy it provides in addressing basic writing errors (Meyer et al., 2024). However, relatively few studies have directly compared these approaches, particularly in EFL contexts, to determine their relative effectiveness. This gap informs the rationale for our study.

This debate between traditional human feedback and emerging AI tools logically gives rise to three distinct models for investigation: teacher-led feedback, AI-only feedback, and a hybrid approach combining both. To establish a clear rationale for our experimental design, the following sections will systematically review the established advantages and disadvantages of each of these three modalities.

### Advantages and Disadvantages of Teacher E-Feedback

Teacher e-feedback, grounded in the traditional approach to feedback but delivered through digital platforms, offers several advantages. One of the primary benefits is its ability to provide personalized,

context-rich responses that are tailored to the individual student's needs (Erkan, 2022). Teachers can draw upon their deep understanding of the students' abilities, progress, and specific challenges to offer nuanced feedback that goes beyond mere error correction, often addressing higher-order concerns such as argument structure, coherence, and task response (Noordin & Khojasteh, 2021). Moreover, teacher e-feedback allows for flexibility in the mode of delivery whereby teachers can provide written comments, audio recordings, or even video feedback, which can enhance the clarity and depth of the feedback provided (Nourinezhad et al., 2021). However, despite these advantages, teacher e-feedback is not without its limitations. The process is often time-consuming, requiring significant effort from educators to assess each piece of writing thoroughly and provide detailed comments (Erkan, 2022). This can lead to delays in feedback delivery, which may hinder students' ability to apply the suggestions to their work immediately (Yu et al., 2021). Additionally, the subjective nature of human feedback can sometimes result in inconsistencies, where different teachers might provide varying levels of detail or focus on different aspects of the writing (Nazari et al., 2021).

### **Advantages and Disadvantages of AI Feedback**

AI feedback systems, powered by advancements in natural language processing and machine learning, offer a contrasting approach to traditional teacher feedback. These systems excel at providing immediate, objective responses to student writing, making them particularly effective for addressing surface-level issues such as grammar, spelling, and punctuation (Baidoo-Anu & Ansah, 2023). The immediacy of AI feedback allows students to quickly recognize and correct their errors, which can be particularly beneficial in contexts where timely feedback is critical (Rasul et al., 2023). Moreover, AI systems can provide consistent feedback, eliminating the variability that might arise from human graders (Bewersdorff et al., 2023). However, AI feedback also has significant drawbacks. One of the most notable is its limited ability to address more complex, higher-order writing issues (Malik et al., 2023). AI systems often struggle with understanding context, nuance, and the rhetorical strategies that are essential for effective writing (Washington, 2023). They may provide inaccurate or irrelevant feedback when addressing more sophisticated aspects of writing, such as argumentation or the effective use of tone and style (Aisyi, 2024). Furthermore, AI feedback lacks the empathetic and supportive dimension that human feedback can offer, which can be critical in building student confidence and promoting long-term engagement with writing tasks (Zhai & Wibowo, 2023).

### **Advantages and Disadvantages of Hybrid Feedback**

Hybrid feedback, which combines the strengths of both teacher e-feedback and AI feedback, presents a promising solution that aims to balance the limitations of each method (Özkanal & Gezen, 2023). By integrating AI's rapid, detailed responses with the personalized, context-rich insights from teachers, hybrid feedback can provide a more comprehensive approach to writing instruction (Connolly & Ceallaigh, 2023; Özkanal & Gezen, 2023). For example, AI can quickly identify and correct grammatical errors, allowing teachers to focus their feedback on higher-order concerns such as content development and the logical flow of ideas. This division of labor can make the feedback process more efficient, potentially reducing the time burden on teachers while ensuring that students receive detailed and meaningful feedback. Additionally, the hybrid approach can cater to different learning preferences, providing immediate, actionable feedback alongside more reflective, in-depth comments (Ghorbandordinejad & Kenshinbay, 2024). However, the hybrid approach is not without its challenges. Integrating AI feedback with human feedback can be complex, requiring careful coordination to ensure that the feedback is coherent and not overwhelming for students (Naseer et al., 2024). There is also a risk that students might become overly reliant on AI for surface-level corrections, neglecting the development of self-editing skills that are crucial for long-term writing improvement (Zhai et al., 2021).

In synthesizing the literature, a central pedagogical tension becomes clear: teacher e-feedback offers deep, contextualized guidance but is constrained by time and scalability, while AI feedback provides immediate, surface-level corrections but often fails to address higher-order rhetorical concerns. The hybrid model is theorized as the optimal solution to resolve this trade-off, yet robust empirical evidence from direct, three-way comparisons is scarce. Therefore, the present study was designed to address this gap by systematically evaluating the relative effectiveness of these three distinct feedback approaches within an Iranian EFL context. Accordingly, we pose the following research questions:

1. Is there a significant difference in the overall writing performance of students who received Teacher E-Feedback, AI Feedback, or Hybrid Feedback? If so, in which specific writing components are these differences significant?
2. What are potential students' perceptions and experiences regarding the use of the approach in which they improved their writing the most?

## Materials and Methods

### Research Design

A randomized controlled trial with a pre-test and post-test design was chosen for this study. This design allows for the comparison of different interventions (teacher e-feedback, AI feedback, and a hybrid approach) while controlling for initial differences between the groups using pre-tests. Participants were randomly assigned to the three feedback groups, ensuring that any observed differences in outcomes could be attributed to the type of feedback provided. The pre-test and post-test measurements of writing performance enabled a precise evaluation of the effectiveness of each feedback method (Campbell & Stanley, 1963; Lee, 2022). This design was selected to ensure rigor in comparing the impact of feedback approaches, and it directly addresses the study's aim to evaluate their relative effectiveness.

### Participants

The participants in this study were 88 adult intermediate students of ESL enrolled in IELTS writing courses at two branches of a single language center in Shiraz, Iran. The study included participants aged between 21 and 35 ( $M = 28$ ,  $SD = 4.2$ ) with an intermediate level of English proficiency, as assessed by a placement test conducted by the language center. Intermediate-level students were specifically selected because it was anticipated that they would possess adequate writing skills necessary for the required writing tasks (Karim & Nassaji, 2019). Moreover, intermediate-level students are more prone to making a range of grammatical and lexical errors in their writing (Brown, 2011).

From the six IELTS writing classes offered at these two branches, three classes were randomly selected to participate in the study. Students in these classes were informed about the nature of the study at the beginning of their two-and-a-half-month semester. They were given the option to switch to other classes if they chose not to participate, and some students opted to change classes, resulting in final class sizes ranging from 28 students (Group 3) to 30 students (Groups 1 and 2).

For the experiment, the three selected classes were randomly assigned to one of the three feedback groups:

- Group 1 (Teacher E-Feedback): Received electronic feedback from teachers.
- Group 2 (AI Feedback): Received feedback from an AI-based writing evaluation tool.
- Group 3 (Hybrid Approach): Received a combination of AI feedback followed by teacher e-feedback.

This random assignment ensured that each group had a balanced number of participants, allowing for a fair comparison of the different feedback methods. Notably, all participants provided informed consent by signing a paper consent form, and their participation was voluntary. These groupings mirror the study's three core feedback conditions, thus ensuring precise evaluation across all methods.

## Instruments

### *Oxford Placement Test (OPT)*

The homogeneity of the participants was assessed using the Oxford Placement Test (OPT) to ensure that there were no significant differences in English proficiency levels among the three groups at the beginning of the study. The results are presented in Table 1. As the baseline measure for participant proficiency, the OPT ensured comparability across all groups at the outset of the study.

**Table 1** *Homogeneity Test Results*

	Sum of squares	df	Mean square	F	Sig.
Between Groups	6.117	2	3.059	.105	.901
Within Groups	2483.439	85	29.217		
Total	2489.557	87			

The ANOVA test results indicated that there were differences in the mean scores of the Oxford Placement Test (OPT) among the three groups; however, these differences were not statistically significant ( $F = 0.105$ ,  $p = 0.901$ ). This indicates that the groups were homogenous in terms of their English proficiency levels at the start of the study, as the p-value is greater than the significance level of 0.05.

### *L2 Essays*

In this study, all groups of L2 learners were asked to write a 200-word essay on a randomly selected topic from Cambridge Book 3 for Task 2, both before the intervention (pre-test) and after it (post-test). The essays produced by participants served as a valuable tool for collecting quantitative data on their writing abilities. The use of essays as both pre- and post-test instruments allowed us to assess the evolution of writing performance across the three groups.

### *IELTS Writing Band Descriptors*

The IELTS Writing Band Descriptors were used to assess the writing improvement of participants in this study, both in the pre-test and post-test. The descriptors evaluate writing based on four key components:

- Task Achievement/Response: This criterion measures how well the test-taker addresses the task requirements and develops ideas.
- Coherence and Cohesion: This criterion evaluates the logical organization of information and ideas, as well as the effective use of cohesive devices.
- Lexical Resource: This criterion evaluates the range and accuracy of vocabulary used in the writing.
- Grammatical Range and Accuracy: This criterion measures the variety and accuracy of grammatical structures used in the writing.



### ***Qualitative Reflection Prompts***

To complement the quantitative results, a set of reflection prompts was used to gather rich qualitative data. This method was chosen as it allows participants to express their thoughts and experiences in a detailed and personal manner, providing in-depth insights into their perceptions of the feedback they received (Ramani & Mann, 2016). The specific prompts administered are detailed in Table 2.

**Table 2** *Reflection Prompt Questions*

Question	Description
General Experience	Describe your overall experience with the x approach. How did you feel about receiving x feedback on your writing assignments?
Impact on Writing Improvement	In what ways did the feedback you received help improve your writing? Can you provide specific examples of how the feedback influenced your writing process?
Impact on Writing Anxiety	Did receiving x affect how you feel about writing tasks, such as making you feel more confident or less anxious about writing? If so, how?
Comparison of Feedback Methods	How would you compare this type of feedback to all the other types of feedback you have ever experienced in terms of usefulness and clarity?
Suggestions for Improvement	Do you have any suggestions for improving the feedback method used in the study? Is there anything you would change about the way feedback was provided?

*Note:* Per the study's explanatory sequential design, the placeholder 'x' was defined as the 'Hybrid Feedback approach' after quantitative results identified this intervention as the most effective.

This study employed an explanatory sequential mixed-methods design, a methodological choice that directly addresses why qualitative data were gathered from only one group. In this design, the primary quantitative phase first identifies which intervention is most effective. The subsequent qualitative phase is then purposefully focused on the highest-performing group—in this case, the Hybrid Group—to explain why it was successful from the learners' perspective. Therefore, selecting only this group for reflection was a deliberate design feature, not an oversight. This approach allows for a “deep-dive” investigation into the specific mechanisms and perceptions that drove the most successful pedagogical outcome, providing richer insights than a more surface-level comparison of all three experiences would yield. To operationalize this design with foresight, the reflection prompts (see Table 2) were initially developed using “x” as a placeholder. This maintained the flexibility to tailor the questions precisely to the intervention that proved most impactful once the quantitative data was analyzed, ensuring our qualitative inquiry was both focused and robust.

### **Data Scoring and Rater Reliability**

To ensure objectivity and minimize potential bias, the scoring of all L2 essays was conducted following a strict protocol. All pre-test and post-test essays were anonymized, assigned a random code, and mixed so that raters were blind to the participant's identity, experimental group assignment (Teacher, AI, or Hybrid), and the time of the test (pre-test vs. post-test).

The scoring was performed independently by two experienced EFL writing instructors who were not involved in the study's interventions. Both raters have over 5 years of experience teaching IELTS writing and were trained on the use of the IELTS Writing Band Descriptors for this study. To establish scoring consistency, inter-rater reliability was calculated on 25% of the essays (a total of 44 essays, randomly selected from both pre- and post-tests). A weighted Cohen's Kappa ( $\kappa_w$ ) was calculated to determine the level of agreement between the two raters' scores. The analysis revealed an almost

perfect level of agreement,  $\kappa_w = 0.85$  ( $p < .001$ ), indicating excellent reliability. Discrepancies in scores were resolved through discussion to reach a consensus score.

## Interventions

To ensure instructional consistency and control for variability in teaching style, a single, experienced writing instructor conducted the classes for all three experimental groups. The instructor followed standardized curricula and feedback protocols for each group (as detailed in Tables 3, 4, and 5) to ensure the primary difference between the groups was the feedback modality itself.

### Group 1 (Teacher E-Feedback)

For Group 1, the writing instructor, who was extremely experienced in providing electronic feedback (e-feedback) to students, began giving e-feedback from the second week of the semester. Initially, students learned the basics of paragraph writing, and by the end of the semester, they had progressed to writing essays. Throughout the semester, feedback was progressively tailored to the complexity of the tasks, ensuring students received guidance suited to their current writing skills. The instructor used a screen-capturing tool called Snagit to create feedback videos. These videos, in MP4 format, were sent to students via the Telegram messaging app. To enhance the feedback, the instructor used a Wacom tablet to draw and write directly on the students' papers. Students were required to complete both in-class assignments and homework, for which they were instructed to scan their papers using the CamScanner app to ensure the quality of the images was acceptable. This method allowed for detailed, personalized feedback on students' writing, helping them to understand and correct their mistakes effectively. The personalized feedback, enhanced by a visual component, helped students gain a clear and actionable understanding of how to improve their writing.

Table 3 outlines the feedback guidelines provided by the researchers to the writing instructor for offering supplementary e-feedback to students in Group 1. These guidelines were specifically designed to ensure that feedback covered all aspects of writing and remained consistent throughout the study.

**Table 3** *Feedback Guidelines for Group 1 Instructor*

Feedback type	Guideline for instructor
Content Feedback	Provide suggestions for expanding or clarifying ideas, improving logical flow, and addressing any inconsistencies.
Grammar and Syntax	Identify and correct grammatical errors, provide explanations, and suggest alternative structures where necessary.
Vocabulary	Recommend more contextually appropriate or varied vocabulary to enrich the text's richness. Highlight any repetitive words and suggest synonyms.
Organization and Structure	Evaluate the overall organization of the essay, including the coherence and cohesion of paragraphs. Suggest ways to improve the structure and transitions between ideas.

### Group 2 (AI Feedback)

From the very beginning of the semester, students in Group 2 were introduced to AI-assisted writing tools and educated on their benefits and challenges. The writing instructor provided comprehensive training on how to effectively utilize these AI tools, including instructions on formulating clear and specific prompts to receive the most useful feedback, interpreting and applying AI-generated suggestions for content improvement, and techniques for leveraging AI tools to enhance grammar,

vocabulary, and style. Throughout the semester, students' main source of feedback was AI tools, specifically ChatGPT-4 (OpenAI) and Grammarly (Premium version). ChatGPT-4 provided suggestions for content development, idea organization, and vocabulary enhancement, but sometimes offered generic or context-insensitive feedback, requiring critical evaluation by students. Grammarly delivered real-time corrections for grammar, spelling, and style, improving lexical accuracy, though it focused primarily on surface-level errors and offered limited guidance on argumentative structure. As detailed in Table 4, these prompts were intentionally designed to align with the core assessment criteria of the study (Task Achievement, Coherence/Cohesion, Lexical Resource, and Grammatical Range/Accuracy). The process was as follows: students wrote their initial drafts independently and then used these tools to receive feedback on content, structure, grammar, and style. After incorporating the AI suggestions, students revised their work. They were then instructed to submit both their original draft and the revised version, assisted by AI, to the instructor.

The instructor reviewed both versions to monitor the students' use of AI feedback and provided additional guidance where necessary, especially in cases where the AI tool might have missed nuanced aspects of writing or context-specific requirements (see Appendix). This approach aimed to leverage the benefits of AI feedback while still maintaining the instructor's oversight, ensuring that students developed critical thinking skills in relation to their writing process.

**Table 4** *Group 2 Student and Writing Instructor Feedback Guidelines*

Feedback type	Prompt sample instruction
Content Feedback	"Please review my essay and provide suggestions for expanding or clarifying ideas, improving logical flow, and addressing any inconsistencies."
Grammar and Syntax	"Identify and correct grammatical errors in my essay. Please provide explanations and suggest alternative structures where necessary."
Vocabulary	"Recommend more contextually appropriate or varied vocabulary to enhance the richness of my text. Highlight any repetitive words and suggest synonyms."
Organization and Structure	"Evaluate the overall organization of my essay, including the coherence and cohesion of paragraphs. Suggest ways to improve the structure and transitions between ideas."
Mechanics	"Identify any spelling, punctuation, and formatting errors in my essay to ensure the final draft is polished and professional."

### Group 3 (Hybrid Approach)

For Group 3, a hybrid method was utilized that integrated both AI feedback and instructor e-feedback. At the beginning of the semester, students were introduced to the advantages and challenges associated with AI-assisted writing tools. The writing instructor provided comprehensive training on how to effectively utilize AI tools to receive insightful feedback on content, grammar, and style. Students were taught to critically assess AI-generated recommendations while preserving their individual voices in their writing. This dual approach allowed students to receive immediate, surface-level corrections from AI tools while also benefiting from in-depth, personalized feedback from their instructor.

Throughout the semester, students utilized an approved AI writing assistant, such as ChatGPT, CoPilot, or AI Chat, for all their assignments. The process involved creating an initial draft, obtaining feedback from the AI tool, and revising the work accordingly. Students were required to submit both their original drafts and the revised versions, which were assisted by AI, to the instructor. To ensure the quality of their submissions, students scanned their papers using the CamScanner app. By submitting



both drafts (original and AI-assisted), students received not only AI-driven feedback but also had the opportunity to reflect on how their revisions were influenced by the AI suggestions, thereby fostering a more robust revision process.

In addition to the AI feedback, the instructor provided supplementary e-feedback on the revised submissions. To convey this feedback, the instructor used a screen-capturing tool called Snagit to create video responses, which were shared with students via Telegram. Using a Wacom tablet, the instructor was able to annotate directly on the students' papers, offering detailed and personalized feedback. This dual feedback mechanism—AI-driven and instructor-led—was designed to complement each other and enhance writing improvement.

Table 5 outlines the feedback guidelines provided by the researchers to the writing instructor for offering supplementary e-feedback to students in Group 3 (Hybrid Approach). These guidelines ensured that the instructor's feedback was consistent and focused on higher-order concerns, allowing for a meaningful balance between AI corrections and more nuanced commentary.

**Table 5** *Group 3 Writing Instructor Feedback Guideline*

Feedback type	Description	Example
Content Feedback	While the AI tool provided initial feedback on the content, the instructor could offer more nuanced suggestions for expanding or clarifying ideas, improving logical flow, and addressing any inconsistencies.	"Your introduction is strong, but the second paragraph could benefit from more details about the main argument. Consider adding examples to support your points."
Grammar and Syntax	The AI tool might identify grammatical errors, but the instructor could provide explanations and alternative structures that the AI might not cover.	"In the third sentence, 'their' should be 'there.' Also, consider rephrasing 'due to the fact that' to 'because' for conciseness."
Vocabulary	The AI tool could suggest synonyms, but the instructor could recommend more contextually appropriate or varied vocabulary to enhance the richness of the text.	"Instead of using 'good' repeatedly, try using synonyms like 'beneficial,' 'advantageous,' or 'favorable.'"
Organization and Structure	The instructor could provide feedback on the overall organization of the essay, including the coherence and cohesion of paragraphs, and suggest ways to improve the structure.	"Your essay would benefit from clearer transitions between paragraphs. Try using phrases like 'Furthermore,' 'In addition,' or 'On the other hand' to guide the reader."
Mechanics	While the AI tool can catch most spelling and punctuation errors, the instructor could ensure that the final draft is polished and professional by addressing any remaining issues.	"There are a few spelling errors in the conclusion. Make sure to double-check words like 'accommodate' and 'definitely.'"

## Data Collection

After taking the placement test, participants' homogeneity was assessed using a brief version of the Oxford Placement Test (OPT). A week later, writing instructors administered the writing pre-test to all three groups. The pre-test consisted of a 200-word essay, which was collected and analyzed to gauge participants' initial writing abilities and provide a baseline for comparison. This pre-test allowed us to ensure that the groups were comparable in terms of initial writing proficiency, which is essential for isolating the effects of the different feedback methods.

Following this, the intervention began. Over the two-and-a-half-month semester, the three groups attended regular IELTS writing classes, which were consistent with the other classes offered at the language center. However, as detailed in the intervention section, the mode of feedback received by L2 learners varied throughout the semester. This allowed for a comparative analysis of how different feedback methods (teacher e-feedback, AI feedback, and hybrid feedback) affected their writing improvement. Throughout the semester, these sessions provided consistent exposure to the interventions, ensuring that the outcomes could be attributed to the feedback methods rather than to other variables.

Upon completion of the educational sessions, the writing post-test was administered to all three groups. This post-test included another 200-word essay to assess the participants' progress in writing improvement. The responses were collected and analyzed to compare the writing performance of the three groups. This post-test data served as a direct measure of writing improvement, providing a clear comparison across the three feedback groups.

### Data Analysis

To evaluate the effectiveness of different feedback methods on the writing performance of EFL learners, Analysis of Covariance (ANCOVA) was employed. This method controlled for pre-existing differences and focused on the impact of the feedback interventions on post-test writing performance. ANCOVA was used to analyze overall writing performance, as well as specific components such as Task Achievement/Response, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy. Pairwise comparisons were conducted to identify significant differences between the feedback groups. By controlling for pre-test scores, ANCOVA ensured that any observed differences in post-test performance were attributable to the type of feedback received rather than initial proficiency.

Qualitative data from the reflection prompts were analyzed using thematic analysis by two independent raters to gain insights into students' perceptions and experiences with the feedback methods. This qualitative analysis helped identify key themes related to the effectiveness of feedback and students' perceived improvements in their writing skills. Thematic analysis of the qualitative data complemented the quantitative findings, offering deeper insights into how students engaged with and benefited from the feedback methods.

## Results

**Research Question 1: Is there a significant difference in the overall writing performance of students who received Teacher E-Feedback, AI Feedback, or Hybrid Feedback? If so, in which specific writing components are these differences significant?**

ANCOVA was conducted to examine differences in overall writing performance and its components across the three feedback groups: Teacher E-Feedback (Group 1), AI Feedback (Group 2), and Hybrid Feedback (Group 3).

Table 6 indicates a significant effect of feedback type on overall EFL writing performance,  $F(2, 84) = 14.07$ ,  $p < .001$ , with the Hybrid Feedback group showing the most improvement. The partial eta squared value of 0.251 suggests that 25.1% of the variance in post-test EFL writing scores can be attributed to the type of feedback received. This finding highlights that the Hybrid Feedback approach was the most effective in enhancing overall writing performance compared to both Teacher E-Feedback and AI Feedback alone.

**Table 6** *Tests of Between-Subjects Effects*

Variable		Source	Sum of squares	df	Mean squares	F statistic	p-value	Partial Eta squared
Overall writing performance		Pre-test	1.46	1	1.46	16.77	0.000	0.166
		Group	2.45	2	1.23	14.07	0.000	0.251
		Error	7.31	84	0.09			
EFL writing components	Task Achievement/Response	Pre-test	6.38	1	6.38	29.26	0.000	0.202
		Group	2.97	2	1.48	4.95	0.009	0.105
		Error	25.21	84	0.30			
	Coherence and Cohesion	Pre-test	6.41	1	6.41	21.95	0.000	0.207
		Group	2.45	2	1.23	4.19	0.018	0.091
		Error	24.54	84	0.29			
	Lexical Resource	Pre-test	0.13	1	0.13	0.33	0.566	0.004
		Group	5.99	2	3.00	7.43	0.001	0.150
		Error	33.88	84	0.40			
	Grammatical Range and Accuracy	Pre-test	0.38	1	0.38	1.06	0.306	0.012
		Group	2.60	2	1.30	3.63	0.031	0.080
		Error	30.09	84	0.36			

Note: Effect sizes ( $\eta^2$ ) interpreted per Cohen's guidelines: .01 = small, .06 = medium, .14 = large.

For Task Achievement/Response ( $F(2, 84) = 4.95, p = .009$ ) and Coherence and Cohesion ( $F(2, 84) = 4.19, p = .018$ ), the Hybrid Feedback group demonstrated the highest performance, accounting for 10.5% and 9.1% of the variance in these components, respectively. This suggests that the Hybrid Feedback method was particularly effective in helping students achieve task objectives and organize their ideas more cohesively.

In the Lexical Resource component, significant differences were also noted ( $F(2, 84) = 7.43, p = .001$ ), with AI Feedback outperforming teacher E-Feedback. This indicates that AI Feedback provided substantial benefits in improving students' vocabulary usage, likely due to the specific, data-driven feedback that AI tools can offer. Finally, for Grammatical Range and Accuracy ( $F(2, 84) = 3.63, p = .031$ ), the Hybrid Feedback group showed the most substantial improvement, highlighting the effectiveness of combining AI's precision in grammar checking with the nuanced feedback provided by teachers.

While the test of between-subjects effects confirms that significant differences exist among the feedback groups across various components of writing, it does not specify which particular groups differ from one another. To gain a more detailed understanding of these differences, pairwise comparisons were conducted (see Table 7).

The Hybrid Feedback group consistently outperformed the Teacher E-Feedback and AI Feedback groups in overall writing performance, with significant differences observed in most comparisons ( $p < .001$ ). Notably, the Hybrid Feedback method led to the highest scores in Task Achievement/Response and Coherence and Cohesion. In contrast, AI Feedback significantly enhanced Lexical Resource, outperforming teacher E-Feedback, while Hybrid Feedback was most effective in improving Grammatical

**Table 7** Results of Pairwise Comparisons for EFL Writing and its Components

Variable		Pairwise comparison	Mean difference	SD	p-value
Overall writing performance		Hybrid Feedback vs. Teacher E-Feedback	0.392	0.078	0.000
		Hybrid Feedback vs. AI Feedback	0.313	0.078	0.000
		AI Feedback vs. Teacher E-Feedback	0.079	0.077	0.921
EFL writing components	Task Achievement/ Response	Hybrid Feedback vs. Teacher E-Feedback	0.395	0.144	0.022
		Hybrid Feedback vs. AI Feedback	0.394	0.144	0.023
		AI Feedback vs. Teacher E-Feedback	0.001	0.141	1.000
	Coherence and Cohesion	Hybrid Feedback vs. Teacher E-Feedback	0.412	0.143	0.015
		Hybrid Feedback vs. AI Feedback	0.239	0.143	0.289
		AI Feedback vs. Teacher E-Feedback	0.173	0.140	0.661
	Lexical Resource	Hybrid Feedback vs. Teacher E-Feedback	0.626	0.167	0.001
		Hybrid Feedback vs. AI Feedback	0.190	0.167	0.772
		AI Feedback vs. Teacher E-Feedback	0.436	0.164	0.029
	Grammatical Range and Accuracy	Hybrid Feedback vs. Teacher E-Feedback	0.201	0.158	0.618
		Hybrid Feedback vs. AI Feedback	0.424	0.157	0.026
		AI Feedback vs. Teacher E-Feedback	-0.223	0.155	0.463

Range and Accuracy. These results highlight the overall superiority of the Hybrid Feedback approach across most components, with AI Feedback excelling specifically in vocabulary enhancement.

Based on the results obtained, we now present the findings of research question 2.

### **Research Question 2: What are students' perceptions and experiences regarding the use of the Hybrid Feedback approach (combining AI feedback and Teacher e-feedback) in improving their English writing skills?**

To gather qualitative data on the participants' experiences and perceptions of the feedback methods used in the study, reflection prompts were distributed among all 28 students in Group 3, which demonstrated the best results in terms of writing improvement. However, since participation was voluntary, not all students completed the reflection papers. Therefore, only 11 of 28 students in the Hybrid Feedback group (Group 3) completed the reflection prompts, resulting in a 39% response rate. This limited sample size may introduce non-response bias, as respondents could be more engaged or

confident learners, potentially reducing the representativeness of the qualitative insights. Despite this, the findings provide valuable insights into the effectiveness of the Hybrid Approach. The responses were analyzed using thematic analysis, with two researchers independently coding the data through an iterative process, achieving an inter-rater reliability of  $\kappa = 0.82$ . Discrepancies were resolved through discussion to ensure coding consistency and validity. These results are summarized in Table 8.

**Table 8** *Thematic Analysis of Reflection Prompts*

Theme(s)	Category	Subcategory/Codes	Frequency (%)
General Experience	Positive Experience	– Engaging feedback – Motivating – Comprehensive	9 (82%)
	Mixed Experience	– Initial confusion – Adaptation over time	2 (18%)
Impact on Writing Improvement	Enhanced Writing Skills	– Improved structure – Better vocabulary – Clearer arguments	10 (91%)
	Specific Examples	– Detailed feedback – Immediate corrections	8 (73%)
Impact on Writing Anxiety	Reduced Anxiety	– Increased confidence – Less fear of mistakes	8 (73%)
	Persistent Anxiety	– Initial nervousness – Gradual reduction	3 (27%)
Comparison of Feedback Methods	AI Feedback	– Quick responses – Detailed corrections	9 (82%)
	Teacher E-Feedback	– Personalized feedback – Encouraging comments	7 (64%)
	Hybrid Approach	– Preference for Integrated Approach – Balanced feedback	11 (100%)
Suggestions for Improvement	Feedback Delivery	– More consistent timing – Clearer instructions	5 (45%)
	Additional Support	– More examples – Follow-up sessions	4 (36%)

*Note:* Percentages were calculated by dividing the frequency of each subcategory by the total number of completed reflection papers (11) and multiplying by 100.

The thematic analysis reveals several key insights into students' experiences with the hybrid feedback approach. A significant majority of students (82%) reported a positive experience, expressing that the combination of AI and teacher feedback was both engaging and motivating. One student noted, *"The combination of AI and teacher feedback was really engaging and motivating. It helped me see my writing from different perspectives."* However, 18% mentioned some initial confusion, which subsided as they became accustomed to the feedback method.

In terms of writing improvement, almost all participants (91%) claimed that their writing skills had enhanced. A common sentiment was captured by one student who remarked, *"My essay structure has significantly improved. The AI feedback helped me organize my ideas better, while the teacher's comments helped me refine my arguments."* Specific improvements, such as detailed feedback and immediate corrections, were acknowledged by 73% of students. Furthermore, a notable 73% reported reduced writing anxiety, with one student sharing, *"I feel much more confident in my writing now."*



*Knowing I'll get both AI and teacher feedback makes me less afraid of making mistakes,"* although 27% still felt some initial nervousness, which resolved over time.

While students valued both AI and teacher feedback individually, all participants (100%) believed the Hybrid Approach to be the most beneficial. As one student summarized, *"The AI feedback was great for quick, detailed corrections, but the teacher's personalized comments really helped me understand the 'why' behind the suggestions."* Specifically, 82% appreciated AI feedback for its swift responses and detailed corrections, whereas 64% favored teacher e-feedback for its personalized support and encouragement.

Lastly, some students (45%) suggested improvements regarding feedback delivery, advocating for more consistent timing and clearer instructions. For instance, one student proposed, *"It would be helpful if the timing of AI and teacher feedback could be more synchronized, so we can plan our revisions better."* Additionally, 36% of students requested more examples and follow-up sessions to support their learning better.

## Discussion

The analysis revealed a significant improvement in overall writing performance among the hybrid feedback group ( $\eta^2 = 0.251$ , large effect per Cohen, 1988), which outperformed both the Teacher E-Feedback and AI Feedback groups. This finding aligns with recent research that highlights the benefits of combining human and AI feedback to leverage the strengths of both methods (Ghorbandordinejad & Kenshinbay, 2024). It also supports Chivapong (2020) argument that hybrid intelligence, by integrating human insight with AI, effectively addresses AI's traditional limitations. The observed effect sizes ( $\eta^2$  ranging from 0.080 to 0.251, medium to large per Cohen, 1988) are consistent with prior meta-analyses of feedback interventions in EFL contexts (Karim & Nassaji, 2019), underscoring the practical significance of hybrid feedback for enhancing EFL writing instruction.

In this study, the Hybrid Approach group demonstrated substantial improvement in Grammatical Range and Accuracy. Recent studies, such as the one by Barrot (2023), have emphasized the role of AI tools like Grammarly in helping learners notice and correct errors through immediate feedback. This practical mechanism, where the AI's instant flagging of errors makes them more salient to the learner, likely drove the observed improvement. However, while Grammarly and similar automated writing evaluation (AWE) systems show promise in enhancing writing accuracy, they have limitations, particularly in providing comprehensive metacognitive explanations. This study highlighted that while AWE can promote noticing and support self-directed learning, it often lacks the depth needed for a full understanding of grammatical rules, potentially leading to overcorrection and cognitive overload. The Hybrid Feedback approach, by combining AI's strengths with human feedback, compensates for these limitations. AI provides immediate, detailed corrections that promote error awareness, while human feedback offers the necessary contextual and metacognitive explanations, resulting in a more thorough and balanced learning experience (Erkan, 2022).

In terms of Coherence and Cohesion, the hybrid feedback group outperformed the other groups. Recent studies by Liu et al. (2023) also suggest that hybrid feedback can enhance students' organizational skills by providing scaffolding that bridges the gap between their current abilities and potential development. The qualitative data from our study support this interpretation; students indicated that AI provided effective micro-level corrections on sentence structure, which allowed the human teacher to focus on macro-level guidance regarding argument flow and logical organization. This dual feedback system, combining AI's micro-level corrections with human teachers' macro-level guidance, not only improves form but also content, echoing Connolly and Ceallaigh (2023) conclusions.

Moreover, in the Lexical Resource component, the AI Feedback group performed significantly better than the Teacher E-Feedback group, though the hybrid group achieved the highest overall scores. This outcome is particularly relevant in the Iranian EFL academic writing context, where students often face challenges in expanding their vocabulary and using appropriate lexical resources. Recent research suggests that writing instructors, particularly in EFL settings, may not always emphasize vocabulary development as much as other aspects of writing, such as grammatical accuracy and coherence, which are often prioritized due to curricular constraints and the demands of high-stakes exams like the IELTS (Chen & Flowerdew, 2018; Coxhead et al., 2019). As a result, students may find themselves with gaps in their academic vocabulary, which are critical for success in such exams. AI tools can help address these gaps by providing immediate, targeted feedback on vocabulary usage, thereby supporting students in developing a richer lexical repertoire. The effectiveness of AI feedback in this study may stem from its ability to offer specific, data-driven vocabulary suggestions that complement the broader instructional goals of teachers (Ding & Zou, 2024).

The effectiveness of the Hybrid Feedback approach in improving writing performance was also evident in the students' perceptions. According to the qualitative analysis, students in this group felt more confident and supported due to the combined feedback methods, which helped to improve their writing skills and overall performance. This outcome supports Krashen's (1982) Affective Filter Hypothesis, which suggests that a balanced and supportive feedback system reduces barriers to learning. The integration of AI's prompt and specific feedback with the teacher's nuanced, contextually relevant insights likely contributed to these positive outcomes (Tsao et al., 2021). In our qualitative analysis, students reported feeling more supported by the feedback process, attributing their progress to the balanced guidance provided by the hybrid method (Bakla, 2020).

Furthermore, the Hybrid Feedback group's superior performance across writing components and their positive feedback align with Dual-Process Theory (Evans & Stanovich, 2013), which distinguishes between System 1 (intuitive, quick processing) and System 2 (reflective, deliberate processing). AI feedback addressed immediate concerns (System 1), while teacher feedback offered reflective, deeper guidance (System 2), reducing barriers to learning and fostering a more comprehensive learning experience. Recent research by Wu (2024) also highlights the role of Hybrid Feedback in fostering student autonomy, which may further contribute to reduced anxiety and improved writing outcomes.

Building on this, teacher e-feedback's weaker performance, notably in lexical resources (see Table 7,  $p = .029$ ), likely reflects instructors' workload and inconsistent delivery, which limits tailored vocabulary guidance (Erkan, 2022). This is compounded by evidence that EFL instructors, due to time constraints or training gaps, often prioritize grammatical accuracy over lexical diversity (Fiallos et al., 2025; Wei-Xun & Jia-Ying, 2024). This suggests targeted professional development, such as workshops on lexical scaffolding and feedback design (Zeggelaar et al., 2022), are needed to enhance feedback strategies. AI feedback's lexical advantage ( $p = .001$ ) stems from ChatGPT-4 and Grammarly's corpus-driven suggestions, providing richer word choices than traditional methods (Ding & Zou, 2024).

Notably, AI tools provide immediate, contextualized lexical feedback, which aligns with cognitive theories of vocabulary acquisition that emphasize repeated exposure and contextual reinforcement (Chen et al., 2021), thereby encouraging AI-supported vocabulary teaching. To mitigate the initial nervousness reported by some students, structured interventions such as introductory workshops (Marzuki et al., 2023) and phased feedback integration can be employed to reduce anxiety and build self-efficacy (Hajiyeva, 2024). Students' recommendations (45% sought consistent timing, 36% requested examples) highlight the need for synchronized feedback schedules and follow-up sessions. Systematic coordination between AI and teacher feedback aligns with research advocating for "just-in-time" feedback frameworks to optimize metacognitive engagement (Acosta-Gonzaga & Ramirez-Arellano,

2022). Variability in AI tools (ChatGPT-4's content focus vs. Grammarly's grammar corrections) suggests tailoring tool selection to writing goals, as algorithmic specificity significantly impacts learner outcomes (Khup & Bantugan, 2025), warranting further research.

Situating these findings within a global context reveals how local pedagogical priorities mediate the effectiveness of hybrid feedback. Our results resonate strongly with research from other exam-driven EFL settings, such as China (Chen & Flowerdew, 2018) and East Asia (Liu et al., 2023), where the combined efficiency of AI and the strategic guidance of a teacher aligns well with performance-oriented goals. This study extends the work by demonstrating the hybrid model's significant impact ( $\eta^2 = 0.251$ ), specifically within the high-stakes IELTS preparation context prevalent in Iran. In contrast, the large effect size observed in our study might be less pronounced in educational systems, such as those in some European countries, that prioritize holistic development or creative expression over the formulaic accuracy targeted by high-stakes exams (Barrot, 2023). Furthermore, our qualitative data—which revealed that 64% of students highly valued the teacher's personalized and encouraging feedback—reinforces findings from other collectivist contexts (Han & Hyland, 2015). This challenges the notion of a universally optimal feedback model, suggesting instead that the Hybrid Approach is particularly potent in cultures where the teacher's supportive role is integral to the learning process, effectively complementing the impersonal efficiency of AI.

### Conclusion, Limitations, and Future Research

This study examined the effectiveness of a Hybrid Feedback approach (combining AI tools with teacher e-feedback) in enhancing the writing skills of intermediate EFL learners. However, the findings must be considered in light of several limitations. First, the sample was drawn from a single private language center in Shiraz, Iran, which restricts the generalizability of the results to broader educational contexts and diverse student populations. Second, the qualitative component of the study is constrained by a small sample size ( $n = 11$ ) resulting from a 39% response rate for the reflection prompts. This raises the possibility of non-response bias, meaning the perspectives gathered may primarily reflect those of the most engaged or confident learners. Therefore, while these qualitative insights are valuable for contextualizing the quantitative results, they are illustrative rather than generalizable and must be interpreted with caution.

Despite these limitations, the study's results suggest that the Hybrid Feedback approach holds significant promise. We observed a robust and practically significant improvement in the hybrid group's writing performance ( $\eta^2 = 0.251$ , large effect). This quantitative evidence, supported by the contextual qualitative data, strongly indicates the pedagogical value of this integrated method. Future research should aim to validate these findings through multi-institutional studies across varied educational landscapes to establish broader applicability.

Students expressed a strong preference for the Hybrid Approach, valuing its ability to meet diverse learning needs through the precision and immediacy of AI feedback, combined with teachers' personalized and reflective guidance. However, 45% noted logistical challenges—particularly inconsistent feedback timing and unclear instructions—that could be addressed through structured training workshops for both instructors and students. These workshops, alongside phased scaffolding, could also mitigate initial student nervousness, fostering familiarity with AI tools. Furthermore, teacher training is necessary to address inconsistencies in feedback and limited lexical guidance, thereby enhancing instructors' ability to support diverse writing components. These practical considerations, while not diminishing the method's effectiveness, highlight implementation factors that educators should consider when adopting this approach.

In summary, the Hybrid Feedback approach shows significant potential as a pedagogical tool for enhancing writing skills. Its balanced integration of AI and human feedback was well-received by learners and demonstrated measurable improvements across all writing components. With refinements to feedback delivery, such as synchronized timing and clearer guidelines, and the adoption of AI-supported vocabulary teaching and tailored tool selection, this strategy holds particular promise for diverse EFL writing contexts, as evidenced by its large effect size, which underscores its practical significance for language instruction.

### Pedagogical Implications for Language Teaching and Technology Integration

Table 9 summarizes the key findings from our study, which focuses on the effectiveness of a hybrid feedback model in writing instruction. This model combines both AI and human feedback to enhance student writing outcomes. The table highlights five key implications for educators who are integrating this approach into their teaching practices.

**Table 9** *Hybrid Feedback Model*

Implication	Description
Adopting a Hybrid Feedback Model	Incorporates both AI and human feedback to leverage the strengths of both, improving writing outcomes.
Customizing Feedback Delivery	Tailors feedback through multiple formats (written, audio, screencast) to meet diverse learner needs and promote better engagement.
Providing Clear Instructions and Consistent Feedback	Ensures feedback is delivered consistently with clear guidelines on how to use it, including synchronized AI and teacher feedback schedules and follow-up sessions to address student recommendations for timing and examples, optimizing metacognitive engagement.
Supporting Continued Learning with Scaffolding	Utilizes scaffolding techniques to provide ongoing support, including introductory workshops and phased feedback integration, to reduce student anxiety and foster familiarity with AI tools, enabling students to gradually develop their writing skills over time.
Encouraging Student Autonomy and Reflection	Fosters self-regulated learning by combining immediate feedback from AI with reflective, personalized guidance from teachers, supporting AI-driven vocabulary acquisition for enhanced lexical development.
Enhancing Teacher Professional Development	Provides targeted training, such as workshops on lexical scaffolding and feedback design, to address inconsistencies and limited vocabulary guidance, while tailoring AI tool selection (e.g., ChatGPT-4 for content, Grammarly for grammar) to diverse writing goals.

### References

- Acosta-Gonzaga, E., & Ramirez-Arellano, A. (2022). Scaffolding matters? Investigating its role in motivation, engagement and learning achievements in higher education. *Sustainability*, 14(20), Article 13419. <https://doi.org/10.3390/su142013419>
- Aisyi, R. (2024). *EFL students' attitude on the use of artificial intelligence (AI) in academic writing* [Thesis, UIN Ar-Raniry Banda Aceh].
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. <https://doi.org/10.61969/jai.1337500>
- Bakla, A. (2020). A mixed methods study of feedback modes in L2 writing. *Language Learning & Technology*, 24(1), 107–128. <https://doi.org/10.125/44712>



- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584–607. <https://doi.org/10.1080/09588221.2021.1936071>
- Bewersdorff, A., Seßler, K., Baur, A., Kasneci, E., & Nerdel, C. (2023). Assessing student errors in experimentation using artificial intelligence and large language models: A comparative study with human raters. *Computers and Education: Artificial Intelligence*, 5, Article 100177. <https://doi.org/10.1016/j.caeai.2023.100177>
- Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, 15(1), 83–97. <https://doi.org/10.1177/1362168810383345>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally College Publishing Company.
- Chen, M., & Flowerdew, J. (2018). Introducing data-driven learning to PhD students for research writing purposes: A territory-wide project in Hong Kong. *English for Specific Purposes*, 50, 97–112. <https://doi.org/10.1016/j.esp.2017.11.004>
- Chivapong, B. (2020). Hybrid Intelligence. The strength and challenges of getting the human in the AI Loop, a literature study [Unpublished master's thesis, Universiteit Hasselt].
- Connolly, C., & Ceallaigh, T. Ó. (2023). *Innovating assessment and feedback design in teacher education: Transforming practice*. Taylor & Francis. <https://doi.org/10.4324/9781003430537>
- Coxhead, A., McLaughlin, E., & Reid, A. (2019). The development and application of a specialised word list: The case of fabrication. *Journal of Vocational Education & Training*, 71(2), 175–200. <https://doi.org/10.1080/13636820.2018.1471094>
- Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*, 1–53. <https://doi.org/10.1007/s10639-023-12402-3>
- Erkan, G. (2022). The impact of teacher e-feedback on students' writing: A waste of time or a road to success? *Focus on ELT Journal*, 4(1), 46–59. <https://doi.org/10.14744/felt.2022.4.1.4>
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1), Article 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3), 223–241. <https://doi.org/10.1177/1745691612460685>
- Febriani, T. N. (2022). “Writing is challenging”: Factors contributing to undergraduate students' difficulties in writing English essays. *Erudita: Journal of English Language Teaching*, 2(1), 83–93. <https://doi.org/10.28918/erudita.v2i1.5441>
- Fiallos, B. A. G., Pilla, M. P. A., Molina, P. E. C., & Melo, N. N. N. (2025). An in-depth analysis of current approaches to teaching argumentative writing in English as a Foreign Language (EFL) classrooms: Examining effective strategies and methodologies. *Ciencia Latina Revista Científica Multidisciplinar*, 9(1), 10249–10261. [https://doi.org/10.37811/cl\\_rcm.v9i1.16299](https://doi.org/10.37811/cl_rcm.v9i1.16299)
- Ghorbandordinejad, F., & Kenshinbay, T. (2024). Exploring AI-driven adaptive feedback in the second language writing skills prompt: AI technology in language teaching. *EIKI Journal of Effective Teaching Methods*, 2(3), 64–71. <https://doi.org/10.59652/jetm.v2i3.264>
- Hajiyeva, B. (2024). Language anxiety in ESL learners: Causes, effects, and mitigation strategies. *EuroGlobal Journal of Linguistics and Language Education*, 1(1), 119–133. <https://doi.org/10.69760/pn9wgv05>
- Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing*, 30, 31–44. <https://doi.org/10.1016/j.jslw.2015.08.002>



- Karim, K., & Nassaji, H. (2019). The effects of written corrective feedback: A critical synthesis of past and present research. *Instructed Second Language Acquisition*, 3(1), 28–52. <https://doi.org/10.1558/isla.37949>
- Khup, V. K., & Bantugan, B. (2025). Exploring the impact and ethical implications of integrating AI-powered writing tools in junior high school English instruction: Enhancing creativity, proficiency, and academic outcomes. *International Journal of Research and Innovation in Social Science*, 9(3s), 361–378. <https://doi.org/10.47772/IJRISS.2025.903SEDU0022>
- Lee, S.-H. (2022). Experimental study. In *Scholarly research in music* (pp. 126–135). Routledge.
- Liu, C., Hou, J., Tu, Y.-F., Wang, Y., & Hwang, G.-J. (2023). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 31(9), 5614–5632. <https://doi.org/10.1080/10494820.2021.2012812>
- Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis, A. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, 5, Article 100296. <https://doi.org/10.1016/j.ijedro.2023.100296>
- Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2), Article 2236469. <https://doi.org/10.1080/2331186X.2023.2236469>
- Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*, 6, Article 100199. <https://doi.org/10.1016/j.caeai.2023.100199>
- Naseer, F., Khalid, M. U., Ayub, N., Rasool, A., Abbas, T., & Afzal, M. W. (2024). Automated assessment and feedback in higher education using generative AI. In R. Sharma & A. Bozkurt (Eds.), *Transforming education with generative AI: Prompt engineering and synthetic content creation* (pp. 433–461). IGI Global.
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), Article e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>
- Noordin, N., & Khojasteh, L. (2021). The effects of electronic feedback on medical university students' writing performance. *International Journal of Higher Education*, 10(4), 124–134. <https://doi.org/10.5430/ijhe.v10n4p124>
- Nourinezhad, S., Hadipourfard, E., & Bavali, M. (2021). The effect of audio-visual feedback on writing components and writing performance of medical university students in two different modes of instruction, flipped and traditional. *Cogent Education*, 8(1), Article 1978621. <https://doi.org/10.1080/2331186X.2021.1978621>
- Özkanal, Ü., & Gezen, E. E. (2023). Students' attitudes and perceptions of e-feedback types: Online teacher feedback (OTF), online peer feedback (OPF), and automated writing evaluation feedback (AWE). *Journal of Advanced Education Studies*, 5, 53–79. <https://doi.org/10.48166/ejaes.1343506>
- Ramani, S., & Mann, K. (2016). Introducing medical educators to qualitative study design: Twelve tips from inception to completion. *Medical Teacher*, 38(5), 456–463. <https://doi.org/10.3109/0142159X.2015.1035244>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), 41–56. <https://doi.org/10.37074/jalt.2023.6.1.29>

- Tsao, J.-J., Tseng, W.-T., Hsiao, T.-Y., Wang, C., & Gao, A. X. (2021). Toward a motivation-regulated learner engagement WCF model of L2 writing performance. *Sage Open*, 11(2), 1–13. <https://doi.org/10.1177/21582440211023172>
- Washington, J. (2023). *The impact of generative artificial intelligence on writer's self-efficacy: A critical literature review*. SSRN. <https://doi.org/10.2139/ssrn.4538043>
- Wei-Xun, L., & Jia-Ying, Z. (2024). Impact of AI-driven language learning apps on vocabulary acquisition among English learners. *Research Studies in English Language Teaching and Learning*, 2(1), 1–11. <https://doi.org/10.62583/rseltl.v2i1.32>
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58(1), 87–125. <https://doi.org/10.1177/0735633119830764>
- Wu, Y. (2024). Critical thinking pedagogics design in an era of ChatGPT and other AI tools—shifting from teaching “What” to teaching “Why” and “How”. *Journal of Education and Development*, 8(1), 1–10. <https://doi.org/10.20849/jed.v8i1.1404>
- Wulandari, Y. (2022). Effective feedback to improve students' writing skills. *Educalitra: English Education, Linguistics, and Literature Journal*, 1(1), 10–17. <https://doi.org/10.5281/zenodo.12795486>
- Yamalee, E., & Tangkiengsirisin, S. (2019). Effects of integrated feedback on academic writing achievement. *Arab World English Journal (AWEJ)* 10(3), 250–270. <https://doi.org/10.24093/awej/vol10no3.17>
- Zeggelaar, A., Vermeulen, M., & Jochems, W. (2022). Evaluating effective professional development. *Professional Development in Education*, 48(5), 806–826. <https://doi.org/10.1080/19415257.2020.1744686>
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, Article 100134. <https://doi.org/10.1016/j.caeai.2023.100134>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021(1), Article 8812542. <https://doi.org/10.1155/2021/8812542>

## Appendix

### Sample of Integrated Instructor and AI-Generated Feedback

Topic: The impact of remote work on urban economies.

Remote work is good because people don't have to waste time travelling to the office. Many people can now work from home, so they think they do not need to live in big, expensive cities anymore. They are moving to smaller towns where life is cheaper and maybe better. This is a bad thing for the economy of cities. The economy of cities are dependent on people being there for work and spending money on things like lunch and transport. Also, many businesses like cafes and retail stores will lose a lot of customers. So, cities must adapt to this new situation. They need to find new ways to use the empty office buildings. This is a big problem they must face.

AI revised paragraph:

The proliferation of remote work presents a significant socioeconomic challenge to the traditional structure of major urban centers. As a growing number of professionals are no longer tethered to a physical office, many are relocating from expensive metropolitan hubs to smaller, more affordable towns. This migration precipitates considerable economic strain on cities, whose economies rely heavily on the daily commerce generated by a commuter workforce, including transportation and hospitality sectors. Consequently, ancillary businesses such as cafes and retail stores face a sharp decline in clientele. To remain viable, urban areas must proactively reimagine their purpose. One potential solution involves repurposing vacant commercial real-estate into residential housing, thereby addressing both commercial vacancies and housing shortages in a single strategic move.

#### Feedback and Explanation of Revisions:

"I have revised your paragraph to improve its organization, academic tone, and logical flow, as you requested. Here is a breakdown of the key changes:"

- **Organization and Structure:**
  - The original paragraph stated the problem but the ideas were somewhat repetitive. The revised version follows a clearer logical structure:
    1. **Main Idea:** Introduce the central challenge (remote work's impact).
    2. **Cause/Effect:** Explain the migration of workers and the resulting economic strain.
    3. **Specific Consequence:** Detail the impact on local businesses.
    4. **Proposed Solution:** Offer a concrete example of adaptation (repurposing buildings).
- **Cohesion and Transitions:**

LP

PC

You have to pay attention to the topic sentence which has to be directly related to the topic (the effect of remote work)

LP

PC

You have to group similar ideas together (e.g., travelling to the office and transport)