

Examining the Peer and AI-Generated Feedback for Writing Proficiency and Learner Motivation

Authors:

Mehran Davaribina¹, Maryam Aghdamie²

¹Department of English, Ard. C., Islamic Azad University, Ardabil, Iran.

²Department of English, Ard. C., Islamic Azad University, Ardabil, Iran.

Corresponding Author:

Mehran Davaribina, Department of English, Ard. C., Islamic Azad University, Ardabil, Iran.

Received: 15-March-2026, Accepted: 30-March-2026, Online First: 31-March-2026

ABSTRACT:

This study investigated the comparative effectiveness of peer and AI-generated feedback on the writing skills and motivation of 90 intermediate-level female EFL learners in Iran, aged 17-18. Participants were randomly assigned to either a peer-feedback or an AI-generated feedback group. Using a pretest-posttest design, changes in writing performance and motivation were assessed using a writing task and the Motivated Learning Strategies Questionnaire (MSLQ), while the Feedback Perception Scale measured student perceptions of the feedback they received. The results indicated that the AI-generated feedback group showed a significantly greater improvement in writing performance than the peer-feedback group ($p < .001$). Conversely, no significant difference was found in motivation levels between the two groups following the intervention ($p = .103$). Students also perceived the AI-generated feedback significantly more favorably across dimensions like usefulness and fairness.

Keywords: *AI-generated feedback; EFL learners; Motivated Learning Strategies Questionnaire; Learner Motivation.*

1. INTRODUCTION:

The development of proficient writing skills is a cornerstone of second language (L2) acquisition, with instructional feedback recognized as a critical catalyst in this process. Effective feedback not only enhances the quality of student writing but also fosters learner engagement and motivation, which are crucial for sustained effort and success (Wu et al., 2024). In contemporary pedagogical landscapes, the proliferation of digital tools has given rise to two prominent feedback modalities: traditional peer feedback and technologically advanced, AI-generated feedback. While peer feedback has long been lauded for promoting collaborative learning and critical thinking (Li et al., 2024), the recent surge in AI-driven tools offers distinct advantages of immediacy, personalization, and scalability (Banihashem et al., 2024). This has created a pedagogical crossroads, where educators must understand the comparative efficacy of these two modalities, particularly concerning their dual impact on writing proficiency and learner motivation. The pedagogical value of peer feedback is deeply rooted in social constructivist learning theories, which posit that knowledge is co-constructed through social interaction and dialogue (Aoun et al., 2018). In this paradigm, learners are not passive recipients but active participants who build understanding by externalizing their thoughts and negotiating meaning (Ma et al.,

2024). Peer feedback operationalizes this by creating a Vygotskian Zone of Proximal Development (ZPD), where students can scaffold each other's learning and achieve outcomes they might not reach independently (Hussain et al., 2024). This collaborative process, which fosters deep engagement in writing tasks (Jin et al., 2025), can enhance motivation by fulfilling learners' need for relatedness and autonomy (Eacersall et al., 2024). Furthermore, by evaluating the work of others, students develop metacognitive awareness that informs their own writing practices (Zhan & Teng, 2025). However, this process is not without significant challenges. Its efficacy is highly variable (Shaban et al., 2021), and the asynchronous nature of digital platforms, while allowing for thoughtful responses, can also result in delayed feedback, potentially diminishing its formative impact (Moorhouse & Wong, 2022).

AI-generated feedback represents a significant pedagogical shift (Lampou, 2023), presenting a compelling alternative designed to overcome the limitations of immediacy and consistency. Leveraging natural language processing, AI tools can provide learners with real-time, objective corrections (Porter & Grippa, 2020), aligning with principles of formative assessment and constructivist learning by enabling a dynamic revision cycle (D. Chen et al., 2025). From the perspective of Self-Determination Theory (SDT), AI feedback can directly support the basic psychological

need for *competence* by offering immediate, actionable advice that helps students feel more capable and effective (Guo et al., 2025). Its impartial nature can also lower writing anxiety (L. Wang et al., 2024). However, current AI models, while proficient at identifying surface-level errors, often lack the capacity to evaluate higher-order rhetorical concerns (Z. Wang et al., 2024). Critically, AI lacks the socio-emotional nuance and empathetic understanding of human interaction, which is vital for managing students' emotional responses and maintaining motivation (Z. Wang et al., 2024). There is also the risk of fostering a dependency where students engage in shallow processing rather than deep reflection (Zhai, 2024).

While separate bodies of literature have extensively examined the merits and drawbacks of peer and AI-generated feedback (Huisman, 2017), there remains a conspicuous gap in research that directly compares their relative impacts within a single, controlled pedagogical context. The interplay between the collaborative, social benefits of peer feedback and the immediate, data-driven advantages of AI feedback is not well understood. Critically, how these different feedback channels—especially when mediated by technology (Liu et al., 2016)—influence the affective domain of student motivation alongside tangible improvements in writing skills remains largely underexplored (Guo et al., 2024). This inquiry is particularly critical in English as a Foreign Language (EFL) settings, such as Iran, where motivational factors can be highly influential on learning persistence and outcomes.

To address this multifaceted gap, the present study investigates the differential effects of peer- and AI-generated feedback on the writing skills and motivation of Iranian intermediate-level EFL learners. Through a comparative experimental design, this research aims to provide robust empirical evidence to guide educators and instructional designers. By moving beyond a simple evaluation of which tool is "better," this study seeks to uncover the nuanced ways in which each feedback type supports or hinders learning, a process that can be hindered when feedback is not immediate (Schofield et al., 2025). Furthermore, it explores student perceptions to provide a holistic view of their implementation in a real-world classroom. Accordingly, the following research questions are posed:

1. What is the impact of peer feedback and AI-generated feedback on the writing skills of Iranian intermediate-level EFL learners?
2. What is the impact of peer feedback and AI-generated feedback on the writing motivation of Iranian intermediate-level EFL learners?
3. What are the perceptions of students toward peer- and AI-generated feedback?

2. METHODOLOGY:

2.1 Participants

The participants of this study consisted of 90 Iranian female EFL students. Their age range was between 17 and 18. They were qualified as learners with intermediate proficiency in English, based on their

performance on Oxford Placement Test (Edwards & Lambert, 2007). Hence, most of the participants had a moderate command of English. These students were randomly assigned into two groups: *peer feedback* and *AI-generated feedback*. Consent was acquired from both students and director of the institute prior to the treatment sessions. Every participant was guaranteed anonymity and could leave at any moment free of retribution.

2.2 Instruments and Materials

A series of instruments, as follows, were used to measure students' general and writing proficiency, motivation and perceptions regarding peer versus AI-generated feedback.

- *The Oxford Placement Test* (Edwards & Lambert, 2007), was used to assess their English language proficiency. The test contains 50 multiple choice questions which assess students' knowledge of key grammar and vocabulary from elementary to intermediate levels, following a reading text with 10 graded comprehension questions. The test's reliability and validity were recognized through prior standardized use, ensuring consistent measurement across participants.
- To assess their writing, students were given a writing topic and asked to develop three paragraphs (not fewer than 300 words) about the positive and negative impacts of technology on society. The learners were given 50 minutes to complete the task. Their writings were scored by both researchers based on a standard scoring rubric, based on Jacobs et al.'s (1981) ESL Composition Profile. The interrater reliability, with an intraclass correlation coefficient (ICC) was 0.9 Cronbach Alpha.
- *The Motivated Learning Strategies Questionnaire* (MSLQ), developed by Pintrich and DeGroot (1990), was administered to measure their motivation for learning and performance, as well as their self-efficacy, including the extent to which students are interested in learning and achieving their goals.
- Furthermore, the *Feedback Perception Scale*, developed by Strijbos et al. (2010), was used to measure the learners' perception concerning the feedback provided by peers and generated by AI. This five-point Likert scale comprised 18 items with 5 dimensions (*fairness, usefulness, acceptance, influence, and willingness to improve*). The total scores could range from 18 to 90 points, with higher scores indicating more positive perceptions of the feedback. This scale was employed to analyze and compare the level of positive perception toward peer-feedback versus AI-

generated feedback across its five dimensions.

2.3 Procedure

Prior to the intervention, participants' English proficiency levels were measured to establish a baseline using the Oxford Placement Test. The results indicated that 86.7% of the students scored at the intermediate level, confirming a homogeneous sample in terms of language competence. The proficiency data ensured that group comparisons accounted for initial language ability, minimizing bias in the interpretation of outcomes. Following this assessment, participants were randomly assigned to two experimental groups, and proceeded with the study's primary tasks, over 10 weeks, (1 session each week).

Since the students had no experience with structured peer evaluation, the teacher carried out a 90-minute training session, explaining the feedback rubric, with some practical examples of the procedure. The teacher also gave specific directions concerning how to compose the writings, (e.g., presenting arguments, academic language, etc.), carry out the motivation questionnaire (rating scales, truthfulness in answers), and take the Feedback Perception Scale (e.g., explaining technical words like "fairness" and "usefulness"). Students practiced with examples to confirm comprehension before proceeding.

In the second session, all participants completed a pre-test consisting of a writing task. They were required to develop a three-paragraph essay on the positive and negative impacts of technology on society. They were allowed 50 minutes to complete the writing task, confirming a controlled assessment of their writing under time limitations and scored based on a standard scoring rubric. Also, the Motivated Learning Strategies Questionnaire (MSLQ) was administered to assess motivation levels of the students.

In sessions 3 to 9, both groups were asked to write several texts, based on the pretest structure, and discuss comments from their peers or AI platform as treatment. After completion, the texts were distributed to peers or submitted to artificial intelligence to provide feedback. Peers in the peer feedback group, assessed each other's drafts using structured feedback forms concentrating on: Content (for example, the clarity of the argument, the examples of the issue), Grammar/Mechanics (for example, the correct use of verb tense, punctuation) and Organization (for example, paragraph coherence, thesis-logic alignment). Students formed small discussion groups to share the comments they received, noticing specific suggestions they found helpful or

unclear. The teacher acted as a facilitator during these discussions, encouraging students to discuss clearly, how they planned to apply the feedback to future writing tasks.

In the AI feedback group, the students had a different method to accomplish the same goal of analyzing their writing and feedback. Each session students reviewed the AI-generated comments, and implemented those suggestions to revise and rewrite their writings. The AI-generated feedback group received feedback generated by ChatGPT 4.0. The AI bot was prompted to provide learners with feedback similar to peer feedback group, i.e., on Content, Organization, Grammar and Mechanics.

After 7 sessions of treatment, in session 10, a post-test was administered to both groups, mirroring the structure of the pre-test with almost similar writing topic. It was the writing task and MSLQ. The goal was to measure any improvement in writing and motivation compared to pretest. Both groups were then given the Feedback Perception Scale to measure their opinions of the feedback they got with respect to acceptance, fairness, value, willingness to improve, and effect. In particular, it was given at the post-test phase to get a whole understanding of the level of positive perception of students regarding the type of feedback the learners would like to receive. This allowed for direct assessment of the influence and power of the two feedback channels.

2.4 Data analysis

Quantitative data from pre- and post-tests (writing ability scores and MSLQ scores) and the perceived feedback scale were analyzed using independent samples t-tests to compare the two groups (peer feedback and AI-generated feedback) on writing performance, motivational changes as well as their perceptions towards these types of feedback.

3. RESULTS AND FINDINGS:

Normality tests, conducted as an important assumption for parametric tests, revealed that the distribution of the scores on writing and motivation in the pre- and post-test conditions were normal. Consequently, to compare the effect of peer feedback and AI-generated feedback on students' writing and motivation, independent samples t-tests were used.

Table 1 shows the descriptive statistics for the scores on Oxford Placement test. As the table clearly indicates majority of the participants were in the intermediate level of proficiency (86.7%).

Table 1. Descriptive Statistics for the Scores on Oxford Placement Test

		Frequency	Percent	Valid Percent	Cumulative Percent	Mean	Std. Deviation	Variance
Valid	pre-intermediate (scored 30-50)	12	13.2	13.3	13.3			
	Intermediate (Scored 50-70)	78	86.7	86.7	100.0			

	Total	90	100	100.0		2.8667	.34184	.117
Total		90	100.0					

As mentioned previously, a writing assessment was performed as pretest in this study. Table 2 indicates descriptive statistics for the scores on this test. The mean for experimental group 1 (peer feedback) in pretest is 11.4667. The mean for experimental group 2 (AI feedback) in pretest is 12.04.

Table 2. Descriptive Statistics for Writing Pretest Scores

Group		Pretest Score	Writing
Experimental 1 peer feedback	Mean	11.4667	
	N	45	
	Std. Deviation	2.07364	
Experimental 2 AI feedback	Mean	12.0444	
	N	45	
	Std. Deviation	2.23562	
Total	Mean	11.7556	
	N	90	
	Std. Deviation	2.16360	

In order to compare the writing performance of the two groups and see whether any significant difference has appeared as a result of the intervention, an independent samples t-test was performed. Table three shows the result.

Table 3. Independent Samples T-test for Writing Pretest Scores

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pretest Writing Score Teacher	Equal variances assumed	.290	.591	-1.271	88	.207	-.57778	.45456	-1.48111	.32556
	Equal variances not assumed			-1.271	87.507	.207	-.57778	.45456	-1.48118	.32563

It can be clearly seen that there is not any significant difference in writing scores between the two experimental groups in the pre-test. ($p > 0.05$). Following the writing test, as mentioned previously, the MSLQ test was administered by the researcher in order to examine students' motivation in the two groups. Table 4 indicates the result.

Table 4. Descriptive Statistics for Motivation Pretest Scores

	N	Minimum	Maximum	Mean	Std. Deviation
Total Pretest Motivation	90	162.00	256.00	224.4667	22.76860
Valid N (listwise)	90				

Similar to the writing test, an independent samples t-test was performed in order to compare the two groups' performance on the motivation pretest. The results appear in Table 5 below.

Table 5. Independent Samples T-test for Motivation Pretest Scores

			Levene's Test for Equality of Variances							t-test for Equality of Means	

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total Pretest Motivation	Equal variances assumed	.000	1.000	.000	88	1.000	.00000	4.82724	-9.59312	9.59312
	Equal variances not assumed			.000	88.000	1.000	.00000	4.82724	-9.59312	9.59312

It can be seen that there is not any significant difference in motivation between the two experimental groups in the pre-test. ($p > .001$). Once the interventions had finished, the students in both groups sat for the writing posttest. Table 6 shows the descriptive statistics for these scores.

Table 6. Descriptive Statistics for Writing Post-test Scores

Group	Post test Writing Score
Experimental 1 peer feedback	Mean
	12.9556
	N
	45
Experimental 2 AI feedback	Std. Deviation
	1.39733
	Mean
	14.9111
Total	N
	45
	Std. Deviation
	1.50487
	Mean
	13.9333
	N
	90
	Std. Deviation
	1.74691

To specify whether these two types of treatment influenced the participants' writing ability differently, an independent-samples t-test was run. Table 7 shows the outcome.

Table 7. Independent samples T-test for the Scores on Writing Post-test

		Levene's Test for Equality of Variances								
		t-test for Equality of Means								
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total posttest Writing Score	Equal variances assumed	26.876	.000	-41.215	88	.000	-5.33333	.12940	-5.59050	-5.07617
	Equal variances not assumed			-41.215	69.111	.000	-5.33333	.12940	-5.59148	-5.07519

The results in table 7 showed a significant difference in writing assessment performance between the two groups in the post-test ($p < .001$). Following the writing posttest the students sat for MSLQ posttest. Table 8 shows the result.

Table 8. Descriptive Statistics for Motivation Post-test Scores

	N	Minimum	Maximum	Mean	Std. Deviation
Total Post test Motivation	90	162.00	283.00	237.6000	23.71545
Valid N (listwise)	90				

In order to find out whether the intervention differently influenced the learners' motivation, an independent samples t-test was run on the post test scores. Table 9 shows the result.

Table 9. Independent samples T-test for Motivation Post-test

		Levene's Test for Equality of Variances								
		t-test for Equality of Means								

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total Pretest Motivation	Equal variances assumed	.000	1.000	.000	88	1.000	.00000	4.82724	-9.59312	9.59312
	Equal variances not assumed			.000	88.000	1.000	.00000	4.82724	-9.59312	9.59312
Total Post test Motivation	Equal variances assumed	2.721	.103	-5.900	88	.000	-25.11111	4.25618	-33.56937	-16.65285
	Equal variances not assumed			-5.900	83.123	.000	-25.11111	4.25618	-33.57630	-16.64593

The independent-samples t-test revealed no significant difference in motivation between the two groups in the post-test ($p > 0.05$). Table 10 below indicates the descriptive statistics for participants' scores on the Feedback Perception Scale, which measured five dimensions (acceptance, fairness, usefulness, willingness to improve, and influence) for both feedback groups.

Table 10. Descriptive Statistics for Peer vs. AI-generated Feedback Perception Scores

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Peer Feedback	45	46.00	61.00	56.49	4.92
AI Feedback	45	68.00	87.00	78.42	5.70

The results in table 10, show that participants who received AI-generated feedback ($M = 78.42$, $SD = 5.70$) reported significantly more positive perceptions than those who received peer feedback ($M = 56.49$, $SD = 4.92$), with both groups showing relatively small variability in their responses.

Table 11. Independent Samples Test for Peer vs. AI-generated Feedback Perception Scores

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total PFB	Equal variances assumed	94.305	.000	29.324	88	.000	21.4888	.73280	20.0326	22.9451
	Equal variances not assumed			29.324	84.601	.000	21.4888	.73280	20.0120	22.9657
Total AIFB	Equal variances assumed	91.429	.000	13.628	88	.000	11.5777	.84953	9.88951	13.2660
	Equal variances not assumed			13.628	67.087	.000	11.5777	.84953	9.86566	13.2899

The results of the independent samples t-test indicated a statistically significant difference between perception of peer feedback and AI feedback ($p < .05$). Based on the attained means, this difference indicates a significantly more positive perception of AI feedback compared to peer feedback.

4. DISCUSSION:

This study investigated the impact of peer-generated and AI-generated feedback on writing skills and motivation among intermediate EFL learners. It also explored student perceptions of these two feedback modalities.

The first research question addressed the impact of feedback type on writing skills. The quantitative results

revealed a statistically significant advantage for the AI feedback group, which demonstrated greater improvement ($M=14.91$) compared to the peer feedback group ($M=12.96$). This finding, supported by a large effect size ($t=-41.215$, $p<.001$), suggests that AI-generated feedback is more effective for improving technical writing proficiency.

This outcome aligns with Cognitive Load Theory (Sweller, 2011), which posits that learning is optimized when extraneous cognitive processing is minimized. The immediate, consistent, and structured nature of AI feedback, particularly on grammatical and syntactic errors, appears to reduce the cognitive burden on learners during the revision process. This allows them to focus more effectively on correcting specific, rule-

based errors, a conclusion supported by recent studies (Edwards, 2007). Conversely, the limited improvement in the peer feedback group likely stems from the inconsistent quality of the feedback provided. As intermediate learners themselves, peers may miss errors or provide incorrect suggestions due to their own developing language proficiency (Xiaoning Liu et al., 2024). Despite these limitations for quantitative gains, peer feedback offers unique benefits. It can foster a collaborative atmosphere (Connolly et al., 2024) and enhance metacognitive skills by exposing students to diverse perspectives (Wei et al., 1998). Furthermore, it is particularly helpful in addressing subjective content and style gaps that benefit from diverse viewpoints (Hussain et al., 2024).

The second research question explored the effect of feedback type on writing motivation. Interestingly, despite the significant differences in writing outcomes, neither feedback modality produced a statistically significant change in student motivation from pre-test to post-test ($p > 0.05$). Although the AI group reported a slightly higher mean motivation score, the difference was not significant.

This finding suggests that the type of feedback alone may not be a primary driver of motivation in this context. This supports Dörnyei's (1998) argument that motivation is a complex construct influenced by numerous contextual factors beyond the specific form of corrective feedback. The result, however, contrasts with literature that has found positive correlations between constructive feedback and motivation (Meşe & Sevilen, 2021). This discrepancy could be attributed to several factors, including the duration of the intervention or the limitations of the Motivated Strategies for Learning Questionnaire (MSLQ) in capturing nuanced motivational shifts related to feedback. It is plausible that while AI's efficiency can be initially motivating, its impersonal nature may temper long-term motivational gains, a point echoed by Zhi and Wang (2024).

The impact of AI-generated feedback on student motivation presents a complex picture. On the one hand, its immediate and personalized nature can be highly motivating for some students (Xiaohong Liu et al., 2024), offering correct feedback without fear of judgment (Jin et al., 2024). On the other hand, constant reliance on it may impede the development of self-regulation and metacognitive awareness.

The present study found no significant impact of peer feedback on student motivation as measured by the MSLQ, a finding that contrasts with some existing literature (Weng et al., 2023). This inconsistency might be attributed to factors such as feedback quality, insufficient student training, and the overall classroom environment (Z. Chen et al., 2025). A supportive learning atmosphere is crucial for effective peer feedback to enhance motivation and foster a sense of community and responsibility (Goagoses et al., 2025). The third research question examined student perceptions of the two feedback types. Here, the results were definitive: students rated AI-generated feedback

significantly higher across all dimensions ($M = 78.42$ vs. $M = 56.49$, $p < 0.05$). Learners perceived the AI feedback as more useful, fair, consistent, and impartial. This strong preference can be understood through the lens of the Technology Acceptance Model, which suggests that perceived usefulness and ease of use are key drivers of technology adoption.

This quantitative preference for AI highlights what Wei and Liu (2024) describe as the "efficiency-empathy divide." Students value the efficiency, objectivity, and accuracy of AI for tangible skill improvement. They appreciated that AI feedback could be received without the fear of judgment that can sometimes accompany peer evaluation (Jin et al., 2024).

The benefits of AI-mediated feedback stem from its immediacy, objectivity, and its focus on easily correctable surface-level errors (Ding & Zou, 2024; Vera et al., 2024), which is highly beneficial for addressing grammatical and stylistic issues promptly. However, it is crucial to acknowledge the limitations of both feedback types. AI-generated systems often primarily focus on superficial errors, potentially neglecting deeper content issues and discouraging self-reflection (Zhai et al., 2025). Students exhibit a wide variety of attitudes toward AI; many value its speed and objectivity, while others find it too impersonal and may disregard its suggestions due to a lack of contextual understanding (Banihashem et al., 2025). Similarly, the effectiveness of peer feedback is intricate, shaped by peer relationships and classroom environment; its credibility is paramount and depends on a supportive collaborative atmosphere (Goagoses, 2025).

Receiving peer feedback in writing can influence learners' writing development. Peer feedback can nurture a learning atmosphere in which students cooperate with their classmates; therefore, it will help them to gain a deeper understanding of the subject (Connolly et al., 2024). The feedback process enables students to critically analyze their own and peers' written work, which in turn enhances their metacognitive and writing skills (Wei et al., 1998). Additionally, students gain alternative perspectives by considering feedback received from peers. This critique may reveal overlooked deficiencies in their writing (Hyland & Hyland, 2006). This method is particularly helpful in addressing subjective content and style gaps that benefit from diverse perspectives (Hussain et al., 2024). The benefits of AI-mediated feedback stem from its immediacy and objectivity. Furthermore, its focus on easily correctable surface-level errors can lead to immediate score improvements, thereby motivating learners to address their writing mistakes.

Both benefits and challenges exist when promoting AI feedback in educational settings. AI software provides immediate and consistent feedback on grammatical and stylistic errors, which is highly beneficial for educators seeking to address issues promptly. This immediate support benefits students with learning difficulties, fostering comprehensive learning. Thus, it is conducive to full-fledged learning. Additionally, AI feedback

offers an objective assessment method, minimizing bias in evaluations, which is crucial in large classes where personalized human feedback may be limited. AI writing correction tools can also clarify grammar rules and writing style for students. This is crucial for developing students' logical thinking and self-regulation skills (Ding et al., 2024).

AI-generated language feedback systems often primarily focus on superficial errors, neglecting deeper content issues. They may provide information that addresses surface-level errors without encouraging a deeper understanding of explanations, meaning, and context, though this is evolving. At times, AI-generated feedback delivered without human mediation might discourage student self-reflection and creativity. Furthermore, AI feedback may lack the human touch, potentially reducing student motivation. Additionally, accuracy remains a key concern, necessitating instructor oversight. Therefore, for AI feedback to genuinely enhance student learning, it should complement human teaching and require a supportive and thoughtfully redefined pedagogy (Zhai et al., 2025). Ultimately, the effectiveness of peer and AI-generated feedback depends on their careful integration within a comprehensive learning model that supports, rather than replaces, human interaction and teaching (Suo et al., 2024). A balanced approach, combining the strengths of peer feedback and AI while acknowledging their limitations, is likely to be more beneficial for students and foster a more effective learning environment. This involves the careful selection of appropriate pedagogical strategies, training of peer assessors and students using AI tools, and the creation of a classroom environment that encourages positive dialogue and critical thinking (Guo et al., 2024).

5. CONCLUSION:

This study provided insight into the distinct effects of AI-generated and peer feedback on the writing skills and motivation of intermediate EFL learners. The central finding reveals a significant divergence: while AI feedback led to substantially greater improvements in writing performance, neither feedback approach produced a measurable change in student motivation. This outcome underscores the intricate relationship between feedback mechanisms and learning, suggesting that the pathways to skill enhancement and motivation are not always aligned. While AI proves highly effective for technical writing development, its pedagogical implementation must be considered alongside other strategies to foster student engagement. The pedagogical implications of these findings are significant. Educators are encouraged to adopt a synergistic model that leverages the strengths of both feedback types. The speed, consistency, and grammatical precision of AI make it an invaluable tool for addressing surface-level errors, thereby improving writing accuracy. Simultaneously, the collaborative and metacognitive benefits of peer feedback remain crucial for developing higher-order thinking and a supportive

learning community. Given that feedback type alone did not impact motivation, teachers should focus on other motivational drivers, such as allowing student choice in writing topics and cultivating a classroom environment that encourages constructive and reflective dialogue.

However, these conclusions should be considered in light of the study's limitations. The research was conducted with a short intervention period and included only female participants of an intermediate proficiency level. Therefore, these findings may not be generalizable to other demographics or proficiency levels. Future research should address these gaps by including diverse participant groups and employing longitudinal designs, such as a pre-test-post-test-delayed post-test model, to track the long-term effects of feedback. Furthermore, qualitative inquiry is needed to delve deeper into why certain feedback features are effective and to explore how individual learner characteristics, such as prior experience and learning styles, influence feedback preferences. Ultimately, while AI tools offer a powerful new resource, their greatest potential lies in a balanced approach that thoughtfully integrates technological efficiency with the irreplaceable value of human interaction and instruction.

6. Pedagogical Implications:

The study offers insights for teachers to reconsider feedback strategies in language learning. The finding that peer and AI feedback did not significantly influence student motivation suggests that other factors contribute to motivation beyond feedback type. Teachers must apply other techniques to enhance writing skills, like helping students to choose the topics for writing and self-reflection through peer review. The findings suggest AI feedback may offer speed advantages, while peer feedback may be more efficient in certain aspects. The synergistic incorporation of both feedback practices would be an ideal solution. Additionally, the research calls for future studies on individual learning preferences regarding feedback, enabling the customization of strategies to help specific students improve their writing skills.

The significant writing improvement seen with AI feedback implies that educators should welcome a synergistic incorporation of AI-mediated feedback with other types of feedback (teacher feedback and peer feedback) in order to facilitate learners' writing improvement.

7. Limitations and Suggestions for Further Research

The present study had some limitations which should be acknowledged here. Firstly, for practical reasons, the study included only female learners. Additionally, while all participants were of intermediate English proficiency to meet experimental conditions, future studies could include learners from other proficiency levels. Furthermore, the intervention duration was short. Future studies can employ a pre-test-post-test-delayed post-test design.

Future research should investigate the specific characteristics of each feedback type that contribute to improved writing quality, such as whether peer feedback focusing on stylistic detail or AI feedback targeting grammar is more impactful. A qualitative analysis can shed light on these relationships. Furthermore, evaluating student characteristics and feedback types (e.g., prior writing experiences and learning styles) may reveal reasons for individual preferences in feedback. The current study's lack of significant motivational impact underscores the importance of identifying effective conditions for peer feedback in enhancing writing skills and motivation.

REFERENCES:

1. Aoun, C., Vatanasakdakul, S., & Ang, K. (2018). Feedback for thought: examining the influence of feedback constituents on learning experience. *Studies in Higher Education*, 43(1), 72-95.
2. Banihashem, S. K., Bond, M., Bergdahl, N., Khosravi, H., & Noroozi, O. (2025). A systematic mapping review at the intersection of artificial intelligence and self-regulated learning. *International Journal of Educational Technology in Higher Education*, 22(1), 50.
3. Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, 21(1), 23.
4. Chen, D., Jia, C., Zhao, X., Yuan, Z., Luo, X., Bao, J., Zhang, H., Zhang, Z., & Hu, X. (2025). Local adaptation is highly dependent on common garden conditions where seeds were propagated: Evidence from a 7-year study on a dominant alpine meadow species. *Journal of Ecology*, 113(3), 689-700.
5. Chen, Z., Wang, J., Liu, K., Gong, J., Sun, L., Wu, Z., Timofte, R., Zhang, Y., Zhang, J., & Wu, J. (2025). NTIRE 2025 challenge on real-world face restoration: Methods and results. Proceedings of the Computer Vision and Pattern Recognition Conference,
6. Connolly, S. J., Sharma, M., Cohen, A. T., Demchuk, A. M., Członkowska, A., Lindgren, A. G., Molina, C. A., Bereczki, D., Toni, D., & Seiffge, D. J. (2024). Andexanet for factor Xa inhibitor-associated acute intracerebral hemorrhage. *New England Journal of Medicine*, 390(19), 1745-1755.
7. Ding, T., Banerjee, A., Mombaerts, L., Li, Y., Borogovac, T., & Weinstein, J. P. D. I. C. (2024). VERA: Validation and evaluation of retrieval-augmented systems. *arXiv preprint arXiv:2409.03759*.
8. Eacersall, D., Pretorius, L., Smirnov, I., Spray, E., Illingworth, S., Chugh, R., Strydom, S., Stratton-Maher, D., Simmons, J., & Jennings, I. (2024). Navigating ethical challenges in generative AI-enhanced research: The ethical framework for responsible generative AI use. *arXiv preprint arXiv:2501.09021*.
9. Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological methods*, 12(1), 1.
10. Edwards, S. L. (2007). Critical thinking: A two-phase framework. *Nurse education in practice*, 7(5), 303-314.
11. Goagoses, N. (2025). The reasons they name: exploring middle school students' achievement goals. *Educational Research and Evaluation*, 1-20.
12. Goagoses, N., Winschiers-Theophilus, H., Auala, S., Pope, N., Rötönen, E., Itege, H., Suero Montero, C., Suovuo, T., & Sutinen, E. (2025). Teachers and Students Envisioning Mixed Reality Remote Learning: A Qualitative Exploration on Fostering Academic Engagement: N. Goagoses et al. *Technology, Knowledge and Learning*, 30(2), 685-709.
13. Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N. V., Wiest, O., & Zhang, X. (2024). Large language model based multi-agents: A survey of progress and challenges. *arXiv preprint arXiv:2402.01680*.
14. Guo, T., Zhong, K., Zhu, N., Li, C., & Kong, F. (2025). Childhood emotional maltreatment and daily pro-environmental behavior among late adolescents: from a self-determination theory perspective. *The Journal of Positive Psychology*, 1-14.
15. Huisman, G. (2017). Social touch technology: A survey of haptic technology for social touch. *IEEE transactions on haptics*, 10(3), 391-408.
16. Hussain, M., Yang, S., Maqsood, U. S., & Zahid, R. A. (2024). Tapping into the green potential: The power of artificial intelligence adoption in corporate green

- innovation drive. *Business Strategy and the Environment*, 33(5), 4375-4396.
17. Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, 39(2), 83-101.
 18. Jin, P., Takanobu, R., Zhang, W., Cao, X., & Yuan, L. (2024). Chat-univi: Unified visual representation empowers large language models with image and video understanding. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,
 19. Jin, Y., Torrent, D., Rouhani, B. D., He, L., Xiang, Y., Xuan, F.-Z., Gu, Z., Xue, H., Zhu, J., & Wu, Q. (2025). The 2024 phononic crystals roadmap. *Journal of Physics D: Applied Physics*, 58(11), 113001.
 20. Lampou, R. (2023). The integration of artificial intelligence in education: Opportunities and challenges. *Review of Artificial Intelligence in Education*, 4, e15-e15.
 21. Li, X., Jiang, S., Hu, Y., Feng, X., Chen, W., & Ouyang, F. (2024). Investigating the impact of structured knowledge feedback on collaborative academic writing. *Education and Information Technologies*, 29(14), 19005-19033.
 22. Liu, P.-H., Hsu, C.-Y., Hsia, C.-Y., Lee, Y.-H., Huang, Y.-H., Chiou, Y.-Y., Lin, H.-C., & Huo, T.-I. (2016). Surgical resection versus radiofrequency ablation for single hepatocellular carcinoma ≤ 2 cm in a propensity score model. *Annals of surgery*, 263(3), 538-545.
 23. Liu, X., Min, X., Zhai, G., Li, C., Kou, T., Sun, W., Wu, H., Gao, Y., Cao, Y., & Zhang, Z. (2024). NTIRE 2024 quality assessment of AI-generated content challenge. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,
 24. Liu, X., Wu, Z., Li, A., Vasluianu, F.-A., Zhang, Y., Gu, S., Zhang, L., Zhu, C., Timofte, R., & Jin, Z. (2024). NTIRE 2024 challenge on low light image enhancement: Methods and results. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,
 25. Ma, R., Chen, K., Andrews, C. B., Loheide, S. P., Sawyer, A. H., Jiang, X., Briggs, M. A., Cook, P. G., Gorelick, S. M., & Prommer, H. (2024). Methods for quantifying interactions between groundwater and surface water. *Annual Review of Environment and Resources*, 49.
 26. Meşe, E., & Sevilen, Ç. (2021). Factors influencing EFL students' motivation in online learning: A qualitative case study. *Journal of Educational Technology and Online Learning*, 4(1), 11-22.
 27. Moorhouse, B. L., & Wong, K. M. (2022). Reconceptualising teacher education courses for online teaching: The experiences of two teacher educators. In *The Use of Technology in English Medium Education* (pp. 109-121). Springer.
 28. Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, 12(24), 10243.
 29. Schofield, S. R., Fisher, A. J., Ginossar, E., Lyding, J. W., Silver, R., Fei, F., Namboodiri, P., Wyrick, J., Masteghin, M. G., & Cox, D. C. (2025). Roadmap on atomic-scale semiconductor devices. *Nano Futures*, 9(1), 012001.
 30. Shaban, M. S., Müller, C., Mayr-Buro, C., Weiser, H., Meier-Soelch, J., Albert, B. V., Weber, A., Linne, U., Hain, T., & Babayev, I. (2021). Multi-level inhibition of coronavirus replication by chemical ER stress. *Nature communications*, 12(1), 5536.
 31. Suo, J., Yang, B., Mosconi, E., Bogachuk, D., Doherty, T. A., Frohna, K., Kubicki, D. J., Fu, F., Kim, Y., & Er-Raji, O. (2024). Multifunctional sulfonium-based treatment for perovskite solar cells with less than 1% efficiency loss over 4,500-h operational stability tests. *Nature Energy*, 9(2), 172-183.
 32. Wang, L., Zhang, X., Su, H., & Zhu, J. (2024). A comprehensive survey of continual learning: Theory, method and application. *IEEE transactions on pattern analysis and machine intelligence*, 46(8), 5362-5383.
 33. Wang, Z., Zhang, J., Du, H., Niyato, D., Cui, S., Ai, B., Debbah, M., Letaief, K. B., & Poor, H. V. (2024). A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications. *IEEE Communications Surveys & Tutorials*, 26(3), 1560-1605.
 34. Wei, R. P., Liao, C.-M., & Gao, M. (1998). A transmission electron microscopy study of constituent-particle-induced corrosion in 7075-T6 and 2024-T3 aluminum alloys.

Metallurgical and Materials Transactions
A, 29(4), 1153-1160.

35. Weng, Y., Wang, Y., Zhang, M., Wang, X., Sun, Q., Mu, S., Wang, H., Fan, M., & Zhang, Y. (2023). Selectively chemocatalytic hydrogenolysis of cellulose to EG and EtOH over porous SiO₂ supported tungsten catalysts. *Catalysis Today*, 407, 89-95.
36. Wu, X., Xing, C., Ju, Q., Su, S., Zhang, X., & Zhu, T. (2024). Research progress in mining ecological restoration technology. *Journal of Industrial Safety*, 1(1), 100004.
37. Zhai, H., Liu, M., Zhang, X., Li, X., Hu, B., Li, H., Gao, S., Wei, Y., & Sun, W. (2025). Science mapping of root ecology: a bibliometric review covering 2015–2024. *Annals of Botany*, mcaf200.
38. Zhai, X. (2024). Transforming teachers' roles and agencies in the era of generative AI: Perceptions, acceptance, knowledge, and practices. *Journal of Science Education and Technology*, 1-11.
39. Zhan, Y., & Teng, M. F. (2025). Assessing English writing self-efficacy beliefs, self-regulation, and performance through asynchronous computer-mediated feedback and face-to-face peer feedback. *Educational Assessment*, 30(1), 4-20.

3