



Reframing Feedback in EFL Paragraph Writing: Complementary Roles of Teachers and ChatGPT-3.5

Huynh Thi Bich Phuong

Ho Chi Minh University of Banking,
VIETNAM

Tran Thanh Du*

Thu Dau Mot University,
VIETNAM

Nguyen Huu The

Ba Ria - Vung Tau College of Education,
VIETNAM

Article Info

Article history:

Received: March 17, 2026

Revised: May 18, 2026

Accepted: June 20, 2026

Keywords:

Academic writing

AI-generated feedback

Human feedback

Paragraph writing

Abstract

The increasing use of generative artificial intelligence in language education has transformed written feedback practices, yet limited evidence exists regarding how teacher and AI-generated feedback differ in supporting distinct dimensions of second language writing. Most previous studies have emphasized feedback effectiveness rather than examining the complementary functions of human and AI feedback across higher-order and lower-order writing concerns. This study investigated the characteristics and functions of teacher and ChatGPT-3.5 feedback on paragraph writing produced by first-year English as a Foreign Language (EFL) students in Vietnam. Guided by Sociocultural Theory, feedback literacy, and the distinction between higher-order and lower-order concerns, the study employed an exploratory mixed-methods design. A total of 178 feedback instances generated from ten student paragraphs were analyzed using quantitative frequency analysis and qualitative thematic coding. The findings revealed that teacher feedback primarily addressed higher-order concerns, including idea development, organization, coherence, and meaning construction, whereas ChatGPT-3.5 predominantly focused on lower-order concerns such as grammar, vocabulary, and sentence structure. Students also perceived teacher feedback as more supportive for conceptual development, while valuing AI-generated feedback for its immediacy and linguistic accuracy. This study extends current understanding of AI-assisted writing by demonstrating that teacher and AI feedback perform complementary rather than competing instructional roles. The findings provide practical implications for integrating human expertise and generative AI to create balanced feedback practices that support both rhetorical development and linguistic accuracy in EFL writing instruction.

To cite this article: Phuong, H. T. B., Du, T. T., & The, N. H. (2026). Reframing feedback in EFL paragraph writing: complementary roles of teachers and ChatGPT-3.5. *Online Learning in Educational Research*, 6(2), 1-10. <https://doi.org/10.58524/oler.v6i2.1220>

INTRODUCTION

Feedback has long been regarded as a key component of second language (L2) writing development since it assists learners to improve both linguistic accuracy and the quality of their ideas. A lot of research has shown that feedback helps writers identify weaknesses, revise their texts, and gradually develop more effective writing practices over time (Bitchener & Ferris, 2012; Hyland, 2019). Feedback plays an active role in shaping how learners approach writing and revision, not simply a corrective tool. From a sociocultural perspective, writing development is not an isolated cognitive activity but a process shaped by interaction and support. Learners do not develop writing skills independently; instead, they rely on interaction, guidance, and support to extend their current abilities (Han & Xu, 2019; Vygotsky, 1978). Within this framework, feedback can be understood as a form of mediation that supports learners in moving beyond their existing level of performance. It provides not only information about errors but also direction for development.

* Corresponding Author

Tran Thanh Du, Thu Dau Mot University, VIETNAM ✉ dutt@tdmu.edu.vn

Traditionally, research in L2 writing has distinguished between higher-order concerns (HOCs) and lower-order concerns (LOCs). Higher-order concerns include content development, organization, and coherence, while lower-order concerns involve grammar, vocabulary, and sentence structure (Bitchener & Ferris, 2012; Sommers, 1982). This distinction is widely used in both research and pedagogy because it reflects different aspects of writing performance. It also suggests that different types of feedback may be required to address different kinds of problems (Ferris, 2010; Mao & Lee, 2020).

However, not all feedback sources address these dimensions in the same way. Teacher feedback has traditionally been associated with meaning-level support. Teachers are able to interpret students' intentions, identify gaps in reasoning, and provide context-sensitive guidance (Fitriani & Sabarniati, 2019; Han & Xu, 2021; Hyland & Hyland, 2006). In practice, this often involves asking students to clarify their ideas, provide additional examples, or reorganize their arguments. At the same time, teacher feedback is shaped by practical constraints. Large class sizes, limited instructional time, and workload pressures can reduce both the amount and the consistency of feedback that teachers are able to provide (Lee, 2008; Bitchener & Ferris, 2012).

These constraints have contributed to increasing interest in automated forms of feedback. With recent developments in artificial intelligence, tools such as ChatGPT-3.5 are now able to generate immediate responses to student writing (Hady et al., 2026; Yakubu et al., 2025). These systems are particularly effective in identifying grammatical errors, suggesting alternative word choices, and revising sentence structures (Mao, 2023; Rafida et al., 2024; Ranalli & Yamashita, 2022). Their ability to provide instant and consistent feedback makes them attractive in contexts where timely teacher input is difficult to obtain. At the same time, the capabilities of AI-generated feedback are not uniform across all dimensions of writing. While these systems perform well in addressing lower-order concerns, their ability to engage with higher-order aspects such as coherence, unity, and idea development remains limited. In many cases, AI feedback focuses on correcting linguistic errors without addressing whether the ideas themselves are clear, relevant, or logically connected. As a result, revisions based on AI feedback may improve grammatical accuracy but leave deeper issues in the text unresolved. This imbalance raises important questions about the role of different feedback sources in supporting writing development.

One way to approach this issue is through the concept of feedback literacy. Feedback is not inherently effective; its value depends on how learners interpret and use it (Carless & Boud, 2018; Carless, 2022; Naufalin et al., 2026). Students vary in their ability to understand feedback, evaluate its relevance, and apply it in revision. Some learners may focus primarily on surface corrections, while others engage more deeply with meaning-level issues. Feedback literacy highlights the importance of learner engagement in the feedback process and suggests that the impact of feedback cannot be fully understood without considering how it is used. As AI-generated feedback becomes more widely available, the role of feedback literacy becomes even more important (Jannah et al., 2026). Unlike teacher feedback, which often includes explanation and guidance, AI feedback tends to provide corrections without context. This places greater responsibility on learners to interpret and evaluate the feedback they receive. Understanding how students respond to different types of feedback, and how they use them in revision is therefore an important area of investigation (Shi & Aryadoust, 2024).

These dynamics become particularly apparent at the level of paragraph writing. Compared to extended academic texts, a paragraph requires writers to construct a coherent unit of meaning within a limited space, requiring them to generate relevant ideas, organize them logically, and express them clearly while maintaining grammatical accuracy. As a result, paragraph writing provides a useful context for examining how feedback operates across different dimensions of writing. Recent studies on AI-assisted writing have shown that AI-generated feedback is particularly effective in identifying grammatical, lexical, and sentence-level problems and can provide immediate support for language accuracy (Choiriyah et al., 2025; Ranalli et al., 2017; Ranalli & Yamashita, 2022; Shi & Aryadoust, 2024). At the same time, teacher feedback continues to demonstrate strengths in addressing higher-order concerns, including idea development, coherence, rhetorical organization, and contextualized guidance (Hyland & Hyland, 2006; Mao & Lee, 2023). Although these studies have advanced the understanding of AI-supported writing instruction, several limitations remain. First, much of the existing research has focused on essay writing, whereas paragraph-level writing has received considerably less attention despite the fact that issues of coherence and idea progression are often

more immediate and visible in shorter texts (Ranalli, 2021). Second, many studies have compared feedback sources primarily in terms of effectiveness or learning outcomes without examining how their functions differ across higher-order and lower-order concerns (Mao & Lee, 2020; Shi & Aryadoust, 2024). Third, although AI tools are increasingly used in higher education, empirical research examining the relationship between teacher feedback and AI-generated feedback in Vietnamese EFL contexts remains limited. These limitations suggest the need for further research examining how teacher and AI-generated feedback differ in addressing various dimensions of paragraph writing, particularly across higher-order and lower-order concerns in EFL contexts.

To address these gaps, this small-scale exploratory study examines human and AI-generated feedback on paragraph writing produced by first-year English majors. To guide the research, the following questions are formulated:

(RQ1) How do teacher and AI-generated feedback differ in their distribution across higher-order and lower-order concerns?

(RQ2) How do the two feedback sources function in addressing students' writing problems?

(RQ3) How do students perceive the usefulness of teacher and AI-generated feedback in paragraph-writing instruction?

METHOD

Research Design

This study employed a small-scale exploratory mixed-methods design to compare human and AI-generated feedback on paragraph writing produced by first-year English majors. The quantitative component focused on examining the frequency and distribution of feedback across higher-order concerns (HOCs) and lower-order concerns (LOCs), while the qualitative component explored the functional characteristics of feedback provided by teachers and ChatGPT-3.5. By combining quantitative and qualitative analyses, the study aimed to provide a more comprehensive understanding of how different feedback sources addressed students' writing problems. Given the limited sample size, the study was exploratory in nature and intended to identify emerging feedback patterns rather than produce broadly generalizable findings.

Participants and Research Context

The study was conducted at a public university in Ho Chi Minh City, Vietnam. A total of 32 first-year English majors completed the classroom writing task, and ten paragraphs were randomly selected for detailed analysis. The selected texts represented a range of writing quality and included various grammatical, organizational, and coherence-related issues commonly observed in early-stage EFL writing. Students' English proficiency levels ranged from pre-intermediate to intermediate based on classroom performance and institutional placement results. Two experienced writing instructors participated as providers of human feedback. Both instructors had more than five years of experience teaching academic writing in EFL contexts and were familiar with providing formative feedback on paragraph writing.

Writing Task and Data Collection

Data collection was conducted during a regular classroom session under controlled conditions. Students were asked to write a short paragraph of approximately 80–100 words in response to the prompt: "What is your dream destination? Describe the place and explain why you want to visit it." Students completed the task individually without using dictionaries, online resources, or AI-based writing tools. After the writing session, the collected paragraphs were reviewed, and ten texts were randomly selected for analysis to maintain a manageable qualitative examination while still capturing a variety of common writing problems found in beginner-level EFL paragraph writing.

Feedback Generation Procedures

Two sources of feedback were collected for each paragraph: teacher feedback and AI-generated feedback. The two instructors independently reviewed and commented on all ten student paragraphs without consulting one another. Although no formal scoring rubric was used, both instructors were asked to provide formative feedback on similar aspects of writing, including content development,

organization, coherence, grammar, vocabulary, and sentence structure. Feedback provided by the two instructors was combined into a single dataset for coding and analysis. Each comment, correction, explanation, or suggestion was treated as one feedback instance.

For AI-generated feedback, each paragraph was submitted individually to ChatGPT-3.5 using the same standardized prompt: "Give your feedback on the following paragraph. Identify grammatical errors, vocabulary issues, sentence structure problems, and suggest improvements for clarity and coherence." The same prompt was consistently applied across all paragraphs to ensure comparability of AI-generated responses. It should be noted that the prompt explicitly emphasized grammatical accuracy, vocabulary, sentence structure, clarity, and coherence. This wording may have influenced the tendency of ChatGPT-3.5 to focus more heavily on lower-order concerns. However, the prompt was intentionally standardized across all paragraphs to maintain consistency in AI-generated feedback. A total of 178 feedback instances were collected, including 92 feedback points from teachers and 86 generated by ChatGPT-3.5.

Student Perceptions

To explore students' perceptions of teacher and AI-generated feedback, brief written reflections were collected after students reviewed both sources of feedback. Students were asked to comment on the usefulness, clarity, and perceived effectiveness of each feedback type in improving their paragraph writing. These reflections were analyzed qualitatively to identify recurring perceptions regarding the strengths and limitations of teacher feedback and AI-generated feedback.

Coding Procedures and Reliability

All feedback instances were first segmented into individual feedback units before coding. The coding framework was developed based on the distinction between higher-order concerns (HOCs) and lower-order concerns (LOCs) commonly used in L2 writing research. Each feedback unit was coded according to two primary categories: (1) focus (higher-order concerns versus lower-order concerns) and (2) analytic dimension (organization and content; grammar and sentence structure; coherence and cohesion). Organization and content, as well as coherence and cohesion, were categorized as higher-order concerns, whereas grammar, vocabulary, and sentence structure were categorized as lower-order concerns.

In addition to the categorical coding, qualitative analysis was conducted to examine the functional characteristics of feedback, including clarification, explanation, correction, and revision guidance provided by teachers and ChatGPT-3.5. To ensure coding reliability, two raters independently coded 25% of the dataset. Inter-rater reliability was calculated using Cohen's kappa and resulted in a coefficient of 0.87, indicating strong agreement. Any disagreements were discussed and resolved before the remaining data were coded. The same coding procedures were subsequently applied to the AI-generated feedback to maintain consistency across both feedback sources.

Data Analysis

Both quantitative and qualitative analyses were conducted. Quantitative analysis focused on comparing the frequency and distribution of feedback categories across teacher and AI-generated feedback. Descriptive statistics were used to identify differences in the emphasis placed on higher-order and lower-order concerns. Qualitative analysis focused on identifying recurring feedback patterns and examining how each feedback source addressed students' meaning-level and form-level writing issues. Representative examples were selected to illustrate differences in the ways teachers and ChatGPT-3.5 responded to problems related to idea development, organization, coherence, grammatical accuracy, and sentence structure.

To provide a clearer overview of the research procedure, the study was conducted in five stages: (1) collection of student paragraphs, (2) generation of teacher and AI-generated feedback, (3) collection of students' written reflections, (4) coding and categorization of feedback instances, and (5) quantitative and qualitative analysis of feedback patterns.

RESULTS AND DISCUSSION

Results

Analysis of the ten student paragraphs written in response to the task “Your dream destination” revealed recurring patterns across both linguistic and discourse-level aspects of writing. Consistent with the study’s analytic framework, these patterns were observed in three main dimensions: organization and content, grammar and sentence structure, and coherence and cohesion. At the level of organization and content, several paragraphs contained underdeveloped or imprecise ideas. Although most students produced identifiable topic sentences, supporting details were often general or insufficiently elaborated. For example, statements such as “China is one of the most famous buildings in the world” and “France is the capital of fashion, so I want to go there” illustrate conceptual inaccuracy and limited explanation. In addition, some paragraphs relied on formulaic sequencing devices (e.g., “First..., Second..., Finally...”) without establishing clear logical relationships between ideas.

At the linguistic level, a range of lower-order errors was observed. These included subject–verb disagreement (e.g., “It gives us knowledge”), incorrect pluralization (e.g., “cherry blossom”), and article misuse (e.g., “in Japanese culture”). Some sentences were structurally unclear or unnatural, such as “People are easy to become addicted.” Issues related to coherence and cohesion were also evident. In several cases, ideas were loosely connected or presented without clear progression, resulting in paragraphs that lacked overall unity. Weak transitions and the absence of concluding sentences were also observed in some responses. A total of 178 feedback instances were analyzed, including 92 from teachers and 86 generated by ChatGPT-3.5. Table 1 presents the overall distribution of feedback across higher-order and lower-order concerns.

Table 1. Distribution of Feedback across Higher-Order and Lower-Order Concerns

Feedback Category		Teacher Feedback (n = 92)	ChatGPT-3.5 Feedback (n = 86)
Higher-order concerns	(organization, content, coherence)	58 (63.0%)	18 (20.9%)
Lower-order concerns	(grammar, vocabulary, sentence structure)	34 (37.0%)	68 (79.1%)
Total		92 (100%)	86 (100%)

The quantitative distribution indicates that teacher feedback focused primarily on higher-order concerns, whereas AI-generated feedback concentrated mainly on lower-order concerns. Teachers frequently commented on idea development, organization, and logical coherence, while ChatGPT-3.5 predominantly addressed grammatical accuracy, vocabulary choice, and sentence structure. Table 2 further presents the distribution of feedback across the three analytic dimensions used in the study.

Table 2. Distribution of Feedback across Analytic Dimensions

Analytic Dimension	Teacher Feedback	ChatGPT-3.5 Feedback
Organization and content	39 (42.4%)	12 (14.0%)
Coherence and cohesion	19 (20.6%)	6 (7.0%)
Grammar and sentence structure	34 (37.0%)	68 (79.0%)
Total	92 (100%)	86 (100%)

The data show clear differences in feedback emphasis between the two sources. Teacher feedback was more evenly distributed across multiple dimensions of writing, whereas AI-generated feedback concentrated heavily on grammar and sentence-level accuracy. Quantitative patterns indicate that teacher feedback was distributed across a wider range of writing dimensions relative to the more narrowly focused AI-generated feedback on sentence-level linguistic accuracy. Teachers’ comments were spread among the issues of organization, content, coherence, and grammar. ChatGPT-3.5, however, focused mainly on grammatical correction and sentence structure. These differences imply that the two feedback sources were different not only in the frequency distribution but also in

the interpretation and priority of students' writing problems. Additional representative examples are presented in Table 3.

Table 3. Comparison of Student Writing Weaknesses and Feedback from Teachers and ChatGPT-3.5

Dimension	Weaknesses from Students' Paragraphs	Teacher Feedback	ChatGPT-3.5 Feedback
Content & Organization (HOC)	<i>"China is one of the most famous buildings in the world."</i>	Identifies conceptual inaccuracy and requests clarification (e.g., specifying a particular building, ...).	No conceptual correction; focuses on grammar only.
	<i>"France is famous so I want to go there."</i>	Requests elaboration and additional contextual information.	Suggests minor rephrasing; no elaboration request.
	<i>Mechanical listing: "First..., Second..., Finally..."</i>	Identifies formulaic structure and suggests alternative organization.	No comment on organization.
	<i>"Japan has delicious food."</i>	Requests supporting examples to develop the idea.	Only minor lexical comments (pluralization).
Grammar & Sentence Structure (LOC)	<i>"It give us knowledge."</i>	Provides explicit grammatical correction.	Provides exact correction: "gives."
	<i>"Cherry blossom in Japan is beautiful."</i>	Notes missing article, may comment on idea vagueness.	Corrects plural form → "Cherry blossoms."
	<i>"People easy to be addicted."</i>	Requests rewrite for clarity AND explanation.	Suggests revised sentence structure: "people can easily become addicted."
	<i>"Traditional Korea foods."</i>	Teacher: "Use 'Korean'; also explain WHY you want to try them."	Fixes to → "traditional Korean foods."
Coherence & Cohesion (HOC)	Ideas unrelated or loosely connected	Identifies lack of alignment between ideas and topic sentence; suggests reordering for flow.	Rarely comments; sometimes suggests "improve coherence" in generic terms.
	Missing concluding sentence	Directly requests a concluding sentence that restates main idea.	Suggests "add conclusion" but without guidance.
	Jumping between ideas	Teacher identifies unclear progression and provides model sequencing.	No detection of coherence problems.
	Weak transitions	Suggests improved use of transitions and logical connectors.	Usually no comment unless transition is grammatically incorrect.

Feedback from humans usually pointed out conceptual misunderstandings and a lack of elaboration. For instance, whenever students provided general statements like "Japan has tasty foods," they were asked by teachers to provide more examples and explain further. If coherence were an issue, feedback from humans would revolve around better sequencing and transitions from one thought to another. On the other hand, ChatGPT-3.5 feedback mainly focused on low-level issues. It always spotted grammatical and lexical mistakes such as subject-verb disagreement, plurality, and sentence formation. Even if there were concept errors and illogical sentences, it would only deal with low-level correction without considering higher-level meanings. To determine student perceptions of feedback effectiveness, students were asked to write reflections based on the feedback that they got from their teachers and the AI tool.

Table 4. Summary of Students' Perceptions of Teacher and AI Feedback

Perception Theme	Teacher Feedback	ChatGPT-3.5 Feedback
Support for idea development	Frequently mentioned	Rarely mentioned
Grammar correction	Moderately mentioned	Frequently mentioned
Clarity of explanation	Frequently mentioned	Occasionally mentioned
Immediate availability	Rarely mentioned	Frequently mentioned
Support for coherence	Frequently mentioned	Limited

Written reflections of students revealed that teacher feedback was overall perceived as more useful for dealing with meaning-level aspects of paragraph writing, particularly idea development, coherence, and overall organization. Teacher comments were often liked by students because they did not just point out weaknesses but also told students why their ideas were unclear, underdeveloped, or not connected well enough. Several reflections revealed that teacher feedback helped students to pinpoint issues they hadn't noticed in their writing, particularly when supporting details were too general or the connection between sentences was not logical. One participant said, "Comments from the teacher show me what is unclear and how to develop the idea better". This implies teacher feedback was perceived as more interpretive and pedagogically supportive, particularly for higher-order concerns.

On the other hand, students appreciated the immediacy, accessibility, and usefulness of ChatGPT-3.5 in identifying grammatical and lexical errors. The feedback generated by AI was often characterized as convenient and easy to understand, especially for students who needed quick help regarding issues at the sentence level, including verb forms, word choice, plural nouns, and sentence structure. Students did, however, also recognize the limitations of AI feedback. Several reflections pointed out that ChatGPT-3.5 rarely helped them determine if their ideas were logical, sufficiently developed, or coherently connected within the paragraph. AI feedback was therefore perceived to be useful for surface-level language correction but not very useful for conceptual problems or improvement of the overall flow of ideas.

Student views align with the findings presented above. The association between teacher feedback and explanations of meanings can be observed. As for AI feedback, it seems to be associated with issues of efficiency, convenience, and accuracy of communication. These perspectives indicate that the two types of feedback could complement each other in terms of teaching paragraph writing. It might be more effective for teachers to influence the rhetorical and conceptual skills of students, while AI feedback could help in editing sentences linguistically.

Discussion

According to the results of this research, there are definite distinctions in terms of what human and AI-generated feedback is like and in terms of the types of aspects that they address. Whereas the former type of feedback was oriented at higher order considerations, including ideas and coherence, the latter one was concerned mostly with lower order issues, such as grammatical, lexical, and structural problems. This research confirmed the orientation of teacher feedback towards meaning. Indeed, instructors often addressed learners' ideas and encouraged them to explain their thinking more thoroughly and coherently (Han & Xu, 2019; Hyland & Hyland, 2006; Ma'arif et al., 2025; Zhang & Hyland, 2022). Based on the findings, teachers appeared to treat paragraph writing as a whole form of communication rather than an assembly of sentences. To understand the results better, it is pertinent to consider the significance of paragraph writing. Unlike long essay writing, which requires students to produce clarity, unity, and progression of ideas within a large volume of content, paragraph writing necessitates this in a very small volume of content. Paragraphs tend to reveal flaws in logical order and the development of concepts in a very clear way.

On the other hand, ChatGPT-3.5 proved particularly efficient in recognizing language errors at the sentence level. The generated feedback made corrections to the mistakes in grammar, modified awkward sentence structures, and enhanced lexical accuracy. This is consistent with prior findings showing that artificial intelligence performs well in resolving surface language issues (Barrot et al., 2023; Mao, 2023; Ranalli & Yamashita, 2022). Nevertheless, the feedback was limited to the linguistic level and hardly ever included interpretation or rhetoric-related aspects. In cases where students had

weakly formulated or logically inconsistent thoughts, the artificial intelligence system tended to focus on language correction instead of meaning-based revision. Additionally, it appears that some features of AI feedback may have been partially shaped by the prompt used in the experiment. Specifically, the explicit mention of the detection of grammatical mistakes, vocabulary errors, issues with sentence structures, and improvements in clarity and coherence could mean that the AI system was prompted towards lower-level concerns. It is important to keep this consideration in mind when interpreting the results of the study.

The second key result is the students' perception of the benefits of teachers' feedback and AI feedback. According to the students' written self-reflections, the teacher's feedback was perceived as being helpful in realizing one's deficiencies in the idea development and cohesion, while AI feedback had mainly an immediate and grammatical nature (Zhai et al., 2021). The results also confirm that different perceptions of feedback are a further argument to believe that teachers' and AI feedback have complementary roles to play in the learning process. One can interpret the findings regarding feedback perceptions in the light of feedback literacy (Syukur et al., 2025). Teachers' feedback typically implied an explanation, clarification, and contextualized assistance, reducing the interpretive load on students. At the same time, AI feedback primarily consisted of corrections without explanations, prompting students to evaluate the value and meaning of the suggestions provided actively. In other words, different feedback sources could impose varying cognitive and interpretive demands on learners.

Additionally, the findings suggest that teacher and AI-generated feedback may provide different forms of support for early-stage EFL paragraph writing (Kohnke et al., 2023; Risdianto et al., 2025). Teacher feedback appears to contribute more strongly to meaning-level guidance and rhetorical organization, while AI-generated feedback offers efficient assistance with sentence-level linguistic accuracy. This complementary relationship may have pedagogical value in writing instruction, particularly in contexts where teachers face time constraints and large class sizes.

LIMITATIONS

This study has several limitations that should be acknowledged. First, the analysis was based on a relatively small sample of paragraph texts produced by first-year English majors at a single Vietnamese university, which may limit the generalizability of the findings. Second, the study examined feedback generated by ChatGPT-3.5 only; other AI systems or newer language models may produce different feedback patterns. Third, the standardized prompt used in the study may have influenced the type of feedback generated by the AI system, particularly its emphasis on lower-order concerns. Finally, the study focused on the characteristics and functions of feedback rather than how students incorporated feedback into revisions or how feedback influenced longer-term writing development. Future research could address these limitations by examining larger datasets, multiple AI tools, and students' revision processes across different educational contexts.

CONCLUSION

This research investigated the attributes and purposes of feedback by teachers and AI on the writing of paragraphs written by students taking their first year English major courses. It was found that the two feedback sources differed in terms of focus. The feedback from teachers centered mainly on higher-order aspects, including idea formation, organization, and coherence, whereas ChatGPT-3.5 emphasized mostly lower-order aspects, specifically grammar, vocabulary, and sentence structure. Although students viewed both feedback types positively, they regarded them differently since they provided different strengths to them.

The implications drawn from the research highlight the need for further studies into the area of AI-assisted writing to demonstrate that there are certain differences between feedback from teachers and feedback from artificial intelligence programs regarding the learning process. Specifically, while feedback from teachers is essential in the process of generating and polishing new ideas, feedback from AI helps resolve language-related problems almost instantly. Instead of considering artificial intelligence tools as an alternative to teachers, the results of the research prove that combining feedback from both sources can result in well-balanced writing support.

AUTHOR CONTRIBUTIONS

Contributions from the authors of this study are outlined based on the CRediT taxonomy. HTBP was involved in the concept formation, methodology development, conducting the investigation, preparing the first draft of the paper, and supervision of the research process. TTD has played roles in data curation, formal analysis, visual representation of data, validation, and reviewing and editing of the manuscript. NHT was involved in providing software support, resources, administration of the project, funding acquisition, and reviewing and editing of the manuscript.

ACKNOWLEDGMENT

We would like to thank every team member of the research group for their valuable contributions to this study.

REFERENCES

- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge. <https://doi.org/10.4324/9780203832400>
- Carless, D. (2022). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. *Active Learning in Higher Education*, 23(2), 143–153. <https://doi.org/10.1177/1469787420945845>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Choiriyah, S., Ramadhan, S., Nugroho, A., Pembangunan, H. R. P., & Muharom, F. (2025). Artificial intelligence-driven learning assessment in faculties of education: An exploratory study. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 6(3), 482–495. <https://doi.org/10.31538/munaddhomah.v6i3.1937>
- Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32(2), 181–201. <https://doi.org/10.1017/S0272263109990490>
- Fitriani, N., & Sabarniati, S. (2019). A feedback investigation of comparing teacher and students' preferences on writing conference in a novice EFL writing class. *Jurnal Ilmiah Peuradeun*, 7(2), 315–330. <https://doi.org/10.26811/peuradeun.v7i2.304>
- Hady, Y., Azkiyah, S. N., Muhib, M., & Maswani, M. (2026). Artificial intelligence-based Arabic language learning: A systematic study of the development and challenges of pedagogical innovation. *Online Learning in Educational Research (OLER)*, 6(1), 49–63. <https://doi.org/10.58524/oler.v6i1.976>
- Han, Y., & Xu, Y. (2019). Student feedback literacy and engagement with feedback: A case study of Chinese undergraduate students. *Teaching in Higher Education*, 26(2), 181–196. <https://doi.org/10.1080/13562517.2019.1648410>
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108635547>
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101. <https://doi.org/10.1017/S0261444806003399>
- Jannah, D. M., Huda, S., Yasin, M., Sanatang, S., Fakhri, M. M., Wahid, S. N., Kaswar, A. B., Soeharto, S., & Amukune, S. (2026). AI literacy, technical skills, and ethical awareness in predicting students' learning performance. *Online Learning in Educational Research (OLER)*, 6(1), 1–12. <https://doi.org/10.58524/oler.v6i1.565>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELJ Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(3), 144–164. <https://doi.org/10.1016/j.jslw.2007.12.001>
- Ma'arif, S., Fathoni, A. M., Suyudi, S., Yudi, U., & El Muna, N. (2025). The interplay of organizational culture, knowledge management system, and learning evaluation in improving English

- language skills: A mixed-method study in Indonesian pesantrens. *Kharisma: Jurnal Administrasi dan Manajemen Pendidikan*, 4(2), 272–286. <https://doi.org/10.59373/kharisma.v4i2.107>
- Mao, Z., & Lee, I. (2020). Feedback scope in written corrective feedback: Analysis of empirical research in L2 contexts. *Assessing Writing*, 45, Article 100469. <https://doi.org/10.1016/j.asw.2020.100469>
- Mao, Z., & Lee, I. (2023). Student engagement with written feedback: Critical issues and way forward. *RELC Journal*, 55(3), 508–528. <https://doi.org/10.1177/00336882221150811>
- Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. Bloomsbury.
- Naufalin, L. R., Indrawati, C. D. S., Indriayu, M., & Noviani, L. (2026). Bridging religiosity and digital competence: The mediating role of digital ethics in higher education. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 7(1), 92–108. <https://doi.org/10.31538/munaddhomah.v7i1.2455>
- Rafida, T., Suwandi, S., & Ananda, R. (2024). EFL students' perception in Indonesia and Taiwan on using artificial intelligence to enhance writing skills. *Jurnal Ilmiah Peuradeun*, 12(3). <https://doi.org/10.26811/peuradeun.v12i3.1520>
- Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, 52, Article 100816. <https://doi.org/10.1016/j.jslw.2021.100816>
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8–25. <https://doi.org/10.1080/01443410.2015.1136407>
- Ranalli, J., & Yamashita, T. (2022). Automated written corrective feedback: Error-correction performance and timing of delivery. *Language Learning & Technology*, 26(1), 1–25. <https://doi.org/10.64152/10125/73465>
- Risdianto, E., Rostika, R. A., Razak, A. Z. A., Kartika, H. A., & Fitria, J. (2025). Unveiling the dual nature of AI in grading: A systematic review of benefits and mitigation strategies for algorithmic bias. *Online Learning in Educational Research (OLER)*, 5(1), 189–216. <https://doi.org/10.58524/oler.v5i1.695>
- Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research: Validity, effects and students' engagement. *ReCALL*, 36(2), 243–262. <https://doi.org/10.1017/S0958344023000265>
- Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, 33(2), 148–156. <https://doi.org/10.2307/357622>
- Syukur, P. A., Fakhri, M. M., Firdaus, F., Putra, K. P., Adiba, F., & Arifiyanti, F. (2025). AI literacy meets ethics: Critical appraisal's mediating role in shaping ethical awareness in higher education. *Online Learning in Educational Research (OLER)*, 5(1), 57–71. <https://doi.org/10.58524/oler.v5i1.508>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Yakubu, M. A., Sain, Z. H., Lawal, U. S., Budiman, S. A., & Permana, D. (2025). Analysis of online academic writing tools usage among postgraduate students in North-west Nigeria. *Kharisma: Jurnal Administrasi dan Manajemen Pendidikan*, 4(1), 1–13. <https://doi.org/10.59373/kharisma.v4i1.69>
- 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, Article 8812542. <https://doi.org/10.1155/2021/8812542>
- Zhang, Z., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51, Article 100586. <https://doi.org/10.1016/j.asw.2021.100586>