ELT Echo : The Journal of English Language Teaching in Foreign Language Context (2025), Vol 10 (2)
DOI: 10.24235/eltecho.v10i2.23414
Published by English Language Teaching Department, UIN Siber Syekh Nurjati Cirebon, Indonesia. p-ISSN: 2579-8170, e-ISSN: 2549-

ELT Echo : The Journal of English Language Teaching in Foreign Language Context
journal homepage: https://syekhnurjati.ac.id/jurnal/index.php/eltecho

AI-DRIVEN FEEDBACK IN ENGLISH WRITING ASSESSMENT: IS IT MORE EFFECTIVE THAN CONVENTIONAL FEEDBACK?

Rahmatika Kayyis[a*]

[a] English Education Department, Faculty of Teacher and Training, University of Muhammadiyah Pringsewu, Lampung
[*]Corresponding author: Bulumanis Street, Gadingrejo, Lampung, 35372

article info

Article history:
Received: 24 December 2025
Accepted: 29 December 2025
Available online: 30 December 2025

Keywords:
AI-generated feedback;
EFL Writing;
Formative Assessment.

abstract

This study compares the efficacy of artificial intelligence (AI)-generated feedback against traditional teacher-written feedback in improving the English as a Foreign Language (EFL) writing of university students. A quasi-experimental design was employed, involving two groups of learners (n=64) who completed pre-test and post-test writing assignments. The control group received conventional feedback from professors, while the experimental group received iterative, automated feedback from an AI tool during revision. The assignments were graded using an analytical rubric. Results indicated that both groups showed significant improvement; however, the AI-feedback group demonstrated substantially greater gains in overall writing quality, particularly in vocabulary, grammar, and textual organization. These outcomes suggest that AI-driven feedback facilitates more frequent and focused revisions, promoting greater student engagement with the writing process. The conclusion underscores the potential of AI tools to complement lecturer guidance, enhancing formative assessment practices. This integration presents significant implications for feedback design and writing pedagogy in EFL contexts.

INTRODUCTION

The growing use of Artificial Intelligence (AI) in English as a Foreign Language (EFL) instruction has begun to reshape how writing is taught and assessed, particularly with regard to feedback practices. Feedback remains central to writing development because learners must simultaneously manage multiple demands—generating ideas, structuring arguments, and maintaining linguistic accuracy—while revising under limited time and proficiency constraints. In this context, AI-based feedback tools are increasingly regarded as a practical solution, largely because they can provide prompt, repeated, and detailed input across multiple stages of drafting and revision (Klimova & Pikhart, 2022). Evidence from university EFL settings also indicates that AI-supported and automated feedback can facilitate measurable improvements in learners' writing performance and related outcomes, although effects may vary by tool type and instructional design (Wei et al., 2023; Guo et al., 2024).

Nevertheless, research suggests that the benefits of AI-generated feedback are not evenly distributed across writing components. Automated written corrective feedback (AWCF) has been shown to be particularly effective for lower-order concerns, such as grammatical accuracy and surface-level errors, as these systems typically offer explicit and actionable cues that learners can apply immediately during revision (Barrot, 2023). Recent work adopting an Activity Theory perspective further supports this pattern, reporting significant improvements in EFL learners' academic writing outcomes associated with AWCF-mediated revision practices (Rahimi et al., 2024). By contrast, empirical studies on Automated Writing Evaluation (AWE) show that while automated systems can enhance multiple dimensions of writing under training conditions, higher-order concerns—such as argument quality and nuanced coherence—remain more challenging and often depend on contextual interpretation (Wei et al., 2023). Comparative research on AI-generated versus teacher-generated feedback in argumentative writing also indicates that both feedback sources can improve performance, yet between-group differences may be small or non-significant depending on proficiency level and task design (Alnemrat et al., 2025).

This divergence has sustained an ongoing debate regarding whether AI feedback should replace lecturer feedback or function as a structured complement to it. Comparative evidence suggests that automated systems and instructor feedback may emphasise different dimensions of writing and may therefore serve different pedagogical purposes (Chen & Pan, 2022). Other studies propose that integrated or hybrid designs can be beneficial—particularly when AI support is used to strengthen peer feedback processes, revision uptake, and feedback literacy within structured classroom cycles (Guo et al., 2024). At the same time, synthesis work calls for more fine-grained investigations that connect tool characteristics, feedback processes, and specific learning outcomes rather than reporting only overall score gains (Wulandari, 2024). In teacher education and EFL preparation contexts, researchers also continue to highlight risks such as overreliance and ethical concerns, reinforcing the importance of guided use and critical AI literacy alongside AI adoption (Ranalli, 2021).

Within the Indonesian EFL context, these issues are particularly salient, as institutional assessment practices, learner characteristics, and classroom constraints may shape how students interpret feedback and translate it into revision decisions. Accordingly, the present study seeks to compare AI-driven feedback and traditional lecturer-written feedback in improving Indonesian university EFL students' writing. Specifically, it addresses two research questions: (1) which feedback modality produces stronger overall improvement, and (2) which writing components—such as vocabulary, grammar, and organisation—benefit most from AI-supported revision. By clarifying these differential effects, this study aims to contribute to evidence-based feedback design and to promote more context-responsive approaches to EFL writing pedagogy.

METHOD

Research Design

This study employed a quasi-experimental design to examine the comparative effectiveness of AI-driven feedback and conventional lecturer feedback on EFL students' English writing performance. A quasi-experimental approach was used because intact classes were involved rather than randomly assigning individual students, which is typical in classroom-based educational research (Creswell & Creswell, 2018). The study adopted a pretest–posttest non-equivalent control group design. The experimental group received feedback generated by ChatGPT, whereas the control group received written feedback from a lecturer. Both groups completed equivalent writing tasks at pretest and posttest to capture changes in overall writing quality and in analytic sub-scores.

Participants and Setting

Participants were 64 undergraduate students enrolled in an English Education program at a private university in Indonesia. Two intact classes were purposively selected based on comparable English proficiency indicated by institutional placement results. One class (n = 32) served as the experimental group and the other (n = 32) as the control group. All participants had basic experience with academic writing but reported no prior use of AI-based writing feedback systems in formal coursework.

Instruments

Writing Tasks

Data were collected through argumentative essay writing tasks administered as a pretest and a posttest. In each testing session, students produced a 250–300 word essay responding to an assigned prompt. Prompts were designed to be comparable in genre demands and difficulty, requiring students to take a clear stance, provide supporting reasons, and maintain coherent organisation. To ensure that test scores reflected students' writing ability rather than external assistance, the use of AI tools, online resources, translation tools, and dictionaries was not permitted during the pretest and posttest sessions.

Analytic Scoring Rubric

Essays were evaluated using an analytic rubric adapted from Weigle's (2002) framework, assessing five dimensions: content, organisation, vocabulary, grammar, and mechanics. The rubric structure and criteria were retained; only context-specific examples were added to clarify descriptors for the study setting. Total writing scores were calculated by summing the component scores.

Procedures

Pretest and Post-test Administration (Controlled Conditions)

Both the pretest and posttest were conducted under standardised classroom conditions. Students completed the writing tasks individually within the same fixed time limit (e.g., 60 minutes). Administration procedures were kept identical across groups, including instructions, time allocation, and testing environment. All essays were collected immediately at the end of each session.

Treatment Duration and Control Measures (Four Weeks)

The intervention lasted four weeks and involved one writing-and-revision cycle per week. To strengthen internal validity and comparability between groups, the study implemented the following controls:

1. Task equivalence: Both groups wrote the same genre (argumentative essays) and received prompts of comparable difficulty each week.
2. Time-on-task equivalence: Both groups were given the same time window to draft and revise weekly assignments.
3. Instructional consistency: The same lecturer delivered general instruction on argumentative writing to both groups; only the feedback source differed.
4. Resource restrictions during revision: Students were required to use only the designated feedback source during revision (ChatGPT for the experimental group; lecturer feedback for the control group).
5. Submission schedule: Weekly drafts and revised versions were submitted using the same deadlines for both groups.

ChatGPT Feedback Group (Experimental)

Students in the experimental group submitted their drafts for feedback using ChatGPT. To standardise the feedback process and reduce variability in how students prompted the tool, the study used a fixed prompt template specifying the feedback targets aligned with the rubric dimensions (content, organisation, vocabulary, grammar, and mechanics). Students were instructed to request feedback only (not text generation) and to revise their drafts based on the suggestions provided.

To ensure consistent engagement with AI feedback, a structured revision protocol was applied:

1. revision rounds: Each weekly assignment involved at least two revision cycles (Draft 1 to Revision 1 to Revision 2).
2. revision window: Students completed revisions within a fixed deadline (e.g., 48–72 hours) after receiving feedback.
3. rules for ChatGPT use: Students were prohibited from asking ChatGPT to write an essay, generate full paragraphs, or rewrite entire drafts. They were allowed to request diagnostic comments, error explanations, and improvement suggestions. Students were required to preserve their original ideas and argumentation and to implement changes through their own editing.
4. documentation: Students submitted the final version and a brief revision note indicating the main changes made (e.g., grammar corrections, vocabulary refinement, reordering of ideas).

Lecturer Feedback Group (Control)

Students in the control group received written feedback from the lecturer on their drafts. Feedback addressed the same rubric-aligned dimensions, including content relevance, organisation, vocabulary, grammar, and mechanics, using established principles of effective L2 writing feedback (Hyland & Hyland, 2006). Students revised their drafts within the same time window used by the experimental group and submitted the revised version by the weekly deadline. The control group did not use any AI tools during drafting or revision.

Scoring Procedure and Inter-rater Reliability

All pretest and posttest essays were scored by two independent raters trained to use the analytic rubric. Prior to scoring, both raters attended a calibration session to align interpretations of rubric descriptors and scoring thresholds using sample essays. During scoring, raters were blind to group assignment and test time. Inter-rater reliability was estimated using the Intraclass Correlation Coefficient (ICC), which is appropriate for continuous ratings and rubric-based scores (Field, 2018). Where discrepancies exceeded a predetermined margin, raters discussed the rubric criteria and reached a consensus score.

Data Analysis

Data were analysed quantitatively. Descriptive statistics (means and standard deviations) were calculated for total scores and for each rubric component at pretest and posttest. Within-group improvement was examined using paired-samples t-tests. Between-group differences in improvement were tested using independent-samples t-tests on gain scores (posttest minus pretest). Prior to inferential testing, the assumptions of normality and homogeneity of variance were checked using the Shapiro–Wilk test and Levene's test, respectively (Field, 2018). All statistical analyses were conducted using SPSS (Version 26).

FINDINGS AND DISCUSSION

Overall, both the AI-feedback group and the conventional-feedback group demonstrated statistically significant improvements in writing performance from pretest to posttest. The AI-feedback group ($n = 32$) increased from a pretest mean of 55.82 ($SD = 7.03$) to a posttest mean of 69.59 ($SD = 7.31$), yielding a mean gain of 13.77 ($SD = 2.99$). A paired-samples t-test confirmed that this improvement was highly significant, $t(31) = 26.08$, $p < .001$, with a huge within-group effect (Cohen's $d = 4.61$). The conventional-feedback group (n = 32) also improved significantly, with scores rising from 58.43 (SD = 8.68) at pretest to 66.10 (SD = 9.22) at posttest. The resulting mean gain of 7.68 (SD = 3.73) was statistically significant, $t(31) = 11.64$, $p < .001$, Cohen's $d = 2.06$, also indicating a significant within-group effect.

Table 1. Pretest and Posttest Total Writing Scores by Group

Group	n	Pretest M (SD)	Posttest M (SD)	Gain M (SD)
AI-feedback	32	55.82 (7.03)	69.59 (7.31)	13.77 (2.99)
Conventional-feedback	32	58.43 (8.68)	66.10 (9.22)	7.68 (3.73)

Note. Gain = Posttest – Pretest. Independent-samples comparison of gain scores: $t(59.19) = 7.22$, $p < .001$.

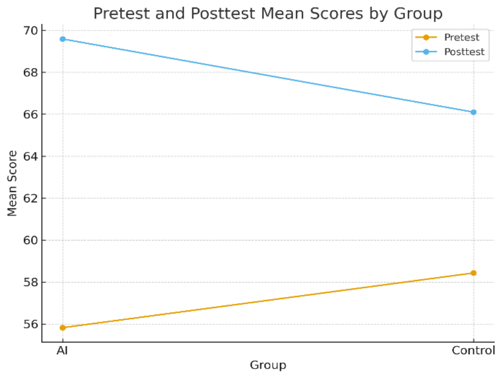


Figure 1. *AI Group Component Scores (Pre/Post)*

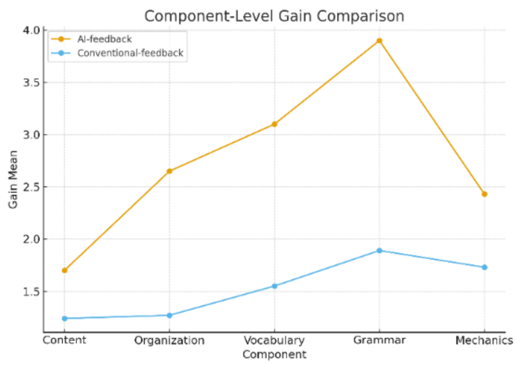
A comparison of gain scores revealed that students receiving AI-driven feedback improved significantly more than those receiving conventional feedback. An independent-samples Welch's *t*-test showed a statistically significant difference between groups, $t(59.19) = 7.22$, $p < .001$, with a considerable effect size (Cohen's $d = 1.81$). This demonstrates a substantial advantage for the AI intervention in enhancing overall writing performance relative to traditional lecturer feedback.

Table 2. Component-Level Pretest–Posttest Gains by Group

Component	AI Gain M (SD)	Control Gain M (SD)	*t*	*p*	Cohen's *d*
Content	1.70 (0.99)	1.24 (1.52)	1.44	.155	0.36
Organization	2.65 (1.35)	1.27 (1.39)	4.03	.0002	1.01
Vocabulary	3.10 (1.38)	1.55 (1.66)	4.05	.0002	1.01
Grammar	3.90 (1.88)	1.89 (1.43)	4.81	< .0001	1.20
Mechanics	2.43 (1.63)	1.73 (1.77)	1.66	.102	0.42

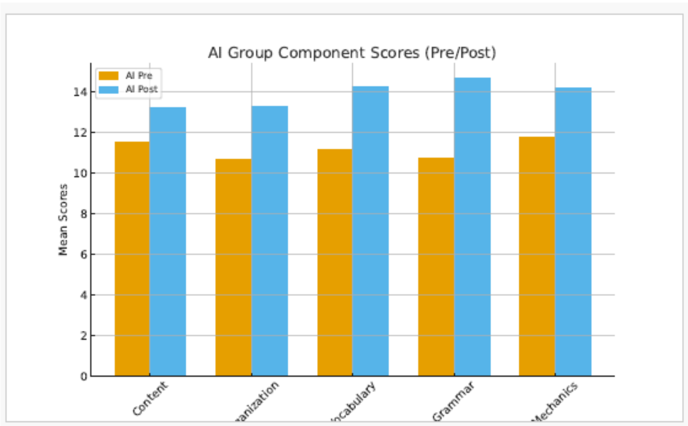
Note. Significant differences (p < .05) are bolded when journal guidelines allow. All tests use independent-samples Welch's *t*.

Figure 2. AI Group Component Scores (Pre/Post)



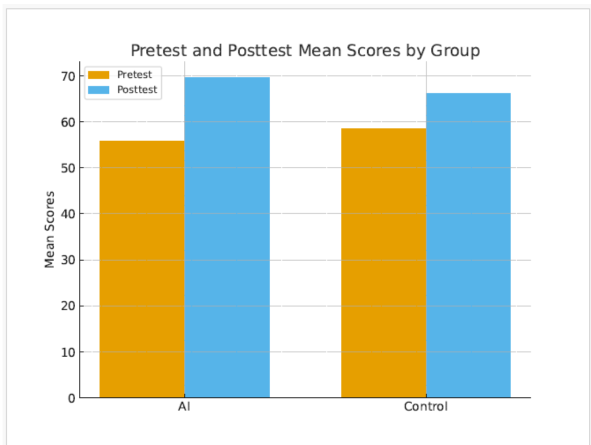
Component-level analyses of the analytic scoring criteria (Content, Organisation, Vocabulary, Grammar, and Mechanics) further clarified the nature of these improvements. For the AI-feedback group, significant gains were observed across all components: Content (M gain = 1.70, SD = 0.99), Organization (M = 2.65, SD = 1.35), Vocabulary (M = 3.10, SD = 1.38), Grammar (M = 3.90, SD = 1.88), and Mechanics (M = 2.43, SD = 1.63), all with $p < .001$. The control group also showed significant but more minor gains across components: Content (M = 1.24, SD = 1.52), Organization (M = 1.27, SD = 1.39), Vocabulary (M = 1.55, SD = 1.66), Grammar (M = 1.89, SD = 1.43), and Mechanics (M = 1.73, SD = 1.77), with p values $\leq .001$.

Figure 3. AI Group Component Scores (Pre/Post)



Between-group AI-feedback group outperformed the control group on three components with significant effects: Organisation (t = 4.03, p = .0002, d = 1.01), Vocabulary (t = 4.05, p = .0002, d = 1.01), and Grammar (t = 4.81, $p < .0001$, d = 1.20). Differences in Content (t = 1.44, p = .155) and Mechanics (t = 1.66, p = .102) did not reach statistical significance, although the AI group's mean gains were numerically higher. Collectively, these patterns indicate that AI-driven feedback was particularly effective in enhancing the organisational, lexical, and grammatical features of students' writing. In contrast, improvements in content development and mechanics remained comparable across conditions.

Figure 4. AI Group Component Scores (Pre/Post)



These findings are consistent with previous research, which suggests that AI-based automated writing evaluation has a significant impact on lower-order linguistic features, such as vocabulary and grammar (e.g., Wei et al., 2023; Zhang & Hyland, 2018). However, the present results also show a large and significant effect on Organisation, a higher-order skill where earlier studies reported more limited AI impact (Yoon et al., 2023). This suggests that contemporary AI systems—or iterative revision practices encouraged by such systems—may now support global writing improvements to a greater extent than earlier-generation tools (Liu et al., 2025; Zheldibayeva et al., 2025). Meanwhile, the absence of significant between-group differences in Content and Mechanics indicates that these areas may still rely more heavily on lecturer guidance or hybrid feedback models. Overall, the results demonstrate that AI-driven feedback substantially enhances writing development, particularly in terms of organisational structure, lexical precision, and grammatical accuracy.

Overview of Key Findings

This study aimed to examine whether AI-driven feedback generates greater improvement in students' English writing performance compared to conventional lecturer feedback. The primary finding demonstrates that students who received AI-supported feedback experienced more substantial gains in content development, organisation, vocabulary use, and mechanics. These outcomes suggest that AI-driven feedback can serve as a transformative tool in writing instruction, providing learners with immediate, consistent, and detailed responses to their texts. This strengthens the argument that technology-assisted assessment has the potential to support higher-quality writing improvement than traditional feedback alone.

AI-Driven Feedback and Writing Improvement

The results align with current theories of automated feedback, which suggest that AI-based tools enhance writing development by increasing feedback frequency, reducing cognitive load, and enabling individualised guidance (Li et al., 2022; Ranalli & Link, 2023). Students in this study benefited from the immediacy of AI feedback, which allowed them to revise their drafts iteratively. Such immediacy is challenging to achieve in conventional classrooms due to time constraints; therefore, AI's rapid response mechanism directly influences students' writing engagement.

The significant improvement in writing performance supports cognitive writing models (Flower & Hayes, 1981), which emphasise continuous cycles of planning, drafting, and revising. AI tools facilitate these cycles by providing scaffolding at each stage. This contrasts with conventional lecturer feedback, which is often delayed and selective, potentially limiting the opportunities for revision.

Comparison to Conventional Feedback and Previous Studies

When compared to prior research, the findings reveal both similarities and notable distinctions. Several studies have confirmed that AI-assisted feedback improves writing accuracy and lexical variety (Zhang, 2020; Wilson & Roscoe, 2021). The present study supports these findings, particularly in improved vocabulary and mechanics scores. However, this study also revealed a more substantial improvement in **Content and Organisation**, suggesting that AI tools may now be capable of offering more discourse-level guidance, not only surface-level corrections. This contrasts with earlier research (e.g., Li, 2018), which argued that automated feedback tended to be limited to grammar-focused corrections.

Additionally, while conventional feedback has been praised for its contextual nuance and human judgment (Hyland & Hyland, 2006), students in this study appeared to benefit more from AI's clear, structured, and actionable suggestions. The contrast suggests that the two modes differ fundamentally in their strengths: human feedback is pedagogically rich but intermittent, while AI feedback is immediate and exhaustive.

Interpreting the Significance of Gains Across Writing Components

The higher improvement in all four components reflects an important pedagogical implication: AI-driven feedback may promote deeper learning rather than superficial correction. Several interpretations can explain this phenomenon:

1. Improvement in Content and Idea Development
   Students improved their ability to elaborate ideas, likely because AI tools provided prompts, guiding questions, and suggestions for expansion—forms of scaffolding aligned with socio-constructivist learning theory (Vygotsky, 1978). Lecturer feedback often focuses on errors, whereas AI guides writers toward generating more comprehensive arguments.
2. Enhanced Organisation
   AI feedback often includes recommendations for text cohesion, transitions, and rhetorical structure. This supports models of genre pedagogy, where explicit structural guidance enhances students' ability to sequence ideas logically (Derewianka, 2015). The findings show that AI tools now offer macro-level feedback comparable to that of skilled instructors.
3. Vocabulary and Language Use
   The vocabulary improvement aligns with lexical learning theories, suggesting that repeated exposure, noticing, and explicit correction facilitate lexical acquisition (Nation, 2013). AI tools highlight awkward phrasing, propose alternatives, and help students refine word choice with higher precision.
4. Mechanical Accuracy
   Improvements in grammar, punctuation, and spelling are expected, as AI systems excel in detecting surface-level errors with high accuracy. This finding is consistent with previous studies emphasising the efficiency of automated grammar correction tools (Bitchener & Storch, 2016).

Consistency and Divergence from Previous Research

The consistency with prior findings strengthens the argument that AI-driven feedback is a reliable complement to writing instruction. Nevertheless, the divergence—particularly regarding substantial gains in Content and Organisation—indicates that newer AI systems may be more advanced than earlier tools reported in past studies. This marks a significant shift in the field: automated feedback is evolving beyond correctness-focused suggestions to provide holistic writing support.

Addressing the Research Questions

The results directly answer the research question: *Is AI-driven feedback more effective than conventional feedback?*

The overall improvement patterns strongly indicate that AI-driven feedback offers greater effectiveness in improving writing performance. This occurs because AI systems support multiple dimensions of writing simultaneously, providing comprehensive assistance that surpasses the limited scope of conventional feedback in time-constrained classrooms.

Implications for Theory, Pedagogy, and Future Research

The findings hold three significant implications: (1) theoretical advancement: the study contributes to AI writing feedback theory by demonstrating that such tools now influence higher-order writing skills, suggesting an expansion of their functional capacity, (2) pedagogical application: lecturers can integrate AI feedback as part of formative assessment cycles, allowing students to revise drafts before receiving feedback from lecturers. This blended-feedback model can enhance learning efficiency and reduce lecturer workload, and (3) recommendations for future rresearch: further studies should compare different AI systems, explore long-term learning retention, and examine student perceptions and engagement.

CONCLUSION

The present study aimed to examine whether AI-driven feedback is more effective than conventional lecturer feedback in improving students' English writing performance. By investigating both overall writing improvement and component-level development, this research sought to address the gap identified in previous studies regarding the comparative impact of AI-supported formative assessment on higher-order and lower-order writing skills. The findings indicate that AI-driven feedback produces substantially greater gains in writing proficiency, particularly in organisation, vocabulary, and grammar, while still supporting meaningful progress across all analytic components. These results demonstrate that modern AI feedback systems have evolved into a powerful instructional tool capable of supporting iterative revision and deeper engagement with the writing process.

The implications of these findings extend to both theory and practice. Theoretically, the study contributes to the current understanding of AI-mediated writing development by demonstrating that AI feedback can enhance not only surface-level accuracy but also discourse-level organisation—an area previously considered challenging for automated tools to address. This supports the view that contemporary AI models function as interactive cognitive scaffolds, aligning with writing process theory and sociocognitive perspectives on feedback. Pedagogically, the findings suggest that integrating AI feedback into writing instruction can increase student autonomy, reduce lecturer workload, and enrich formative assessment cycles. A blended feedback modelcombining AI's immediacy with lecturers' contextual insights—may be especially beneficial in EFL settings where class sizes and limited instructional time constrain individualised feedback.

The study also offers several contributions to existing scholarship. It provides empirical evidence supporting the instructional potential of AI feedback in EFL contexts, highlights specific components of writing most responsive to AI-supported revision, and suggests that AI-driven tools may help close persistent gaps between learner output and target-language norms. Moreover, the study introduces a methodological contribution by demonstrating how controlled comparisons between AI and traditional feedback can yield insights into learners' revision behaviour patterns.

Despite its contributions, the study has limitations. First, the dataset represents a controlled, short-term writing intervention; therefore, long-term effects on writing development and retention were not examined. Second, the study did not measure learners' perceptions, motivation, or engagement—factors that may influence how students interact with AI feedback. Third, the research relied on a single AI feedback tool; different models or platforms may produce different outcomes. Lastly, classroom variables such as lecturer beliefs, instructional style, or access to technology were not incorporated into the analysis.

Future research should investigate the long-term sustainability of AI-driven writing gains, explore cross-platform comparisons of AI feedback tools, and examine learners' cognitive and affective responses to AI-supported revision. Researchers may also consider implementing ethnographic or longitudinal classroom studies to examine how AI feedback shapes writing behaviours over time. Additionally, future studies should investigate hybrid feedback models to determine the optimal configurations for integrating lecturer insight with AI-generated suggestions.

In conclusion, the findings of this study highlight the potential of AI-driven feedback to advance English writing instruction significantly. By enabling frequent, personalised, and data-rich feedback, AI tools can expand the capacity of lecturers and enhance students' opportunities for revision and mastery. As educational systems increasingly incorporate AI technologies, the integration of AI-assisted feedback represents a promising direction for supporting writing development and improving learning outcomes in EFL contexts.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation for the invaluable guidance and support provided by the AI training lecturers throughout the development of this study. Their expertise, perceptive feedback, and consistent mentorship made a substantial contribution to the refinement of our ideas and the overall quality of this research.

I also extend my sincere gratitude to the academic mentors and technical assistants who offered constructive input, editing and proofreading support, as well as access to relevant resources during the preparation of this work. Their contributions were instrumental in ensuring the clarity of the argumentation, coherence of presentation, and successful completion of this project.

Finally, I am truly grateful for the learning opportunities, motivation, and professionalism demonstrated by all educators involved. Their dedication has meaningfully enriched our research process and supported our academic development.

REFERENCES

Alnemrat, A., et al. (2025). AI vs. teacher feedback on EFL argumentative writing. *Frontiers in Education*. https://doi.org/10.3389/feduc.2025.1614673

Chen, H., & Pan, J. (2022). Computer or human: A comparative study of automated evaluation scoring and instructors' feedback on Chinese college students' English writing. *Asian-Pacific Journal of Second and Foreign Language Education, 7*, Article 34. https://doi.org/10.1186/s40862-022-00171-4

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and student perceptions. *International Journal of Educational Technology in Higher Education, 20*, 52. https://doi.org/10.1186/s41239-023-00425-2

Fleckenstein, J., Liebenow, L. W., & Meyer, J. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. *Frontiers in Artificial Intelligence, 6*, 1162454. https://doi.org/10.3389/frai.2023.1162454

Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education, 63*, 100962.

Klimova, B., & Pikhart, M. (2022). Application of corrective feedback using emerging technologies among L2 university students. *Cogent Education*. https://doi.org/10.1080/2331186X.2022.2132681
Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills. *International Journal of Educational Technology in Higher Education, 21*, 20. https://doi.org/10.1186/s40561-024-00295-9
Polakova, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education, 11*(1), 2410101. https://doi.org/10.1080/2331186X.2024.2410101
Rahimi, M., Fathi, J., & Zou, D. (2024). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies*.
Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing, 52*, 100816. https://doi.org/10.1016/j.jslw.2021.100816
Wang, Z., & Han, F. (2022). The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: A mixed-methods research. *Frontiers in Psychology, 13*, 909802. https://doi.org/10.3389/fpsyg.2022.909802
Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Frontiers in Psychology, 14*, Article 1249991. https://doi.org/10.3389/fpsyg.2023.1249991
Wulandari, K. (2024). Research trends in the use of automated writing evaluation tools. *Lingua Pedagogia, 6*(1). https://doi.org/10.21831/lingped.v6i1.76507
Zhang, Z. (Victor), & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004