

ARTIFICIAL INTELLIGENCE-ASSISTED FEEDBACK IN ACADEMIC WRITING: STUDENTS' EXPERIENCES, BENEFITS, AND CHALLENGES IN HIGHER EDUCATION

Juliati^{1*}

Universitas Samudra, Langsa, Indonesia

juliati@unsam.ac.id

Azrul Rizki²

Universitas Samudra, Langsa, Indonesia

azrul.rizki@unsam.ac.id

*Corresponding author: juliati@unsam.ac.id

Article History	Received: July 18, 2024	Revised: September 16, 2024	Accepted: October 21, 2024	Published: November 15, 2024
------------------------	-------------------------	-----------------------------	----------------------------	------------------------------

Abstract

Artificial intelligence-assisted feedback has increasingly influenced academic writing practices in higher education by providing immediate responses to grammar, vocabulary, coherence, clarity, citation style, and organization. However, students' experiences with artificial intelligence-assisted feedback remain complex because these tools can support revision while also creating dependency, uncertainty, and ethical concerns. This study aims to examine students' experiences, perceived benefits, and challenges in using artificial intelligence-assisted feedback for academic writing. The study employed a mixed-method design involving a survey of 104 undergraduate students, semi-structured interviews with 20 students, and analysis of 48 writing drafts revised with artificial intelligence feedback. Quantitative data were analyzed using descriptive statistics, while qualitative data were analyzed thematically. The findings showed that students perceived artificial intelligence-assisted feedback as useful for improving grammar accuracy, sentence clarity, vocabulary choice, organization, and revision confidence. However, students also reported challenges related to inaccurate suggestions, generic feedback, overdependence, difficulty evaluating feedback quality, loss of authorial voice, and uncertainty about academic integrity. Writing draft analysis indicated that students tended to accept surface-level feedback more frequently than higher-order feedback related to argumentation and structure. The study concludes that artificial intelligence-assisted feedback can support academic writing when students use it critically, selectively, and transparently. Pedagogical guidance, feedback literacy, and institutional policy are needed to ensure responsible integration in academic writing instruction.

Keywords: artificial intelligence; academic writing; feedback literacy; student experience; higher education

A. Introduction

Academic writing is a central practice in higher education because students are required to construct arguments, synthesize sources, present evidence, organize ideas, and communicate knowledge in disciplinary contexts. University students are expected not only to produce grammatically correct texts but also to demonstrate clarity, coherence, critical reasoning, rhetorical awareness, and ethical use of sources. However, academic writing remains challenging for many students, particularly those who are still developing academic language competence and genre awareness. They often struggle with sentence accuracy, paragraph coherence, vocabulary precision, argument development, citation integration, and revision strategies.

Feedback plays an essential role in academic writing development. Through feedback, students can identify weaknesses in their writing, understand expectations, revise their drafts, and develop greater awareness of academic discourse. Effective feedback does not merely indicate errors; it guides students to improve meaning, organization, language use, and rhetorical quality. Feedback in second language writing has been widely discussed as a key pedagogical mechanism for developing accuracy, revision behavior, and writing competence (Hyland & Hyland, 2006; Ferris, 1999, 2004; Bitchener & Knoch, 2010; Shintani, Ellis, & Suzuki, 2014). However, providing timely and detailed feedback is often difficult for lecturers because of large class size, limited time, and varied student writing needs.

Artificial intelligence-assisted feedback has emerged as a response to these challenges. In academic writing, artificial intelligence tools can provide automated suggestions related to grammar, spelling, punctuation, vocabulary, sentence clarity, tone, coherence, cohesion, paraphrasing, summarization, and citation format. Earlier forms of automated writing evaluation focused mainly on scoring and surface-level corrective feedback, while recent generative artificial intelligence tools can provide more interactive, explanatory, and text-specific feedback. This development has changed how students approach writing, feedback, and revision (Warschauer & Ware, 2006; Lee, 2020; Ranalli, 2018; Godwin-Jones, 2022).

Nevertheless, artificial intelligence-assisted feedback also raises important concerns. Automated suggestions are not always accurate, context-sensitive, or pedagogically appropriate. Some tools may overemphasize grammar and style while ignoring argument quality, source integration, disciplinary conventions, and authorial intention. Generative artificial intelligence tools may produce fluent suggestions that appear convincing but are not necessarily accurate. Students may

also accept feedback uncritically, resulting in loss of voice, inappropriate paraphrasing, dependency, or unclear authorship boundaries (Chan & Hu, 2023; Chiu, Xia, Zhou, Chai, & Cheng, 2023; Kasneci et al., 2023; Nguyen, Ngo, Hong, Dang, & Nguyen, 2023).

Therefore, the effectiveness of artificial intelligence-assisted feedback depends on students' feedback literacy. Feedback literacy refers to students' ability to understand, interpret, evaluate, and use feedback for learning. In the context of AI-assisted feedback, students need to judge whether suggestions are accurate, relevant, ethical, and aligned with writing goals. They must decide which suggestions to accept, modify, reject, or discuss with lecturers. This requires cognitive, behavioral, affective, and ethical engagement with feedback (Carless & Boud, 2018; Dawson et al., 2019; Molloy, Boud, & Henderson, 2020; Zhang & Hyland, 2022).

Previous studies have shown that students' engagement with automated and teacher feedback is shaped by trust, writing proficiency, feedback clarity, and perceived usefulness. Bai and Hu (2017) found that students respond selectively to fallible automated writing feedback. Zhang and Hyland (2018) showed that students' engagement with teacher and automated feedback involves cognitive, behavioral, and affective dimensions. Koltovskaia (2020) further revealed that automated corrective feedback can support ESL writing, but student engagement determines whether feedback becomes meaningful. Nguyen, Hong, Dang, and Huang (2024) also demonstrate that human-AI collaboration in academic writing varies according to how writers interact with AI suggestions.

Two further strands frame the problem. The first is feedback literacy. Carless and Boud (2018) and Molloy, Boud, and Henderson (2020) argue that students need appreciation of feedback, capacity to make judgements, and strategies for acting on comments before any feedback source can be useful to them; without that capacity, a faster source of feedback simply produces faster copying. The second is the institutional debate about artificial intelligence in education. Zawacki-Richter, Marin, Bond, and Gouverneur (2019) show in a systematic review that most applications of artificial intelligence in higher education have been developed with little reference to pedagogical theory, and Chiu, Xia, Zhou, Chai, and Cheng (2023) map the opportunities and risks that follow. Kasneci et al. (2023) and Nguyen, Ngo, Hong, Dang, and Nguyen (2023) argue that transparency, accountability, and learner autonomy must be built into classroom use rather than added afterwards.

Although research on artificial intelligence in writing is growing, more empirical studies are needed on students' actual experiences and

challenges in using artificial intelligence-assisted feedback. Many discussions focus on whether artificial intelligence tools are beneficial or risky, but fewer studies examine how students interpret feedback, what types of feedback they trust, how they revise their writing, and what difficulties they encounter. This gap is important because students are already using artificial intelligence tools in academic writing, often faster than institutions develop formal guidelines (Strobl et al., 2019; Zawacki-Richter, Marin, Bond, & Gouverneur, 2019).

This study therefore investigates artificial intelligence-assisted feedback in academic writing from students' perspectives. It focuses on students' perceived benefits, challenges, feedback uptake, and ethical concerns. The research questions are: (1) How do students experience artificial intelligence-assisted feedback in academic writing? (2) What benefits do students perceive from using artificial intelligence-assisted feedback? (3) What challenges do students encounter when using artificial intelligence-assisted feedback? (4) How do students use artificial intelligence feedback in revising their academic writing drafts?

B. Method

This study employed a mixed-method research design. The quantitative component used a student survey to identify general patterns of students' experiences, perceived benefits, and challenges in using artificial intelligence-assisted feedback. The qualitative component used semi-structured interviews and writing draft analysis to explore students' experiences more deeply. A mixed-method design was selected because the study aimed to capture both measurable tendencies and contextual interpretations of students' engagement with artificial intelligence feedback.

The participants were 104 undergraduate students enrolled in academic writing courses at a higher education institution. The participants were selected purposively because they had experience using artificial intelligence tools to receive feedback on academic writing drafts. The students came from different study programs, including English education, language studies, social sciences, and teacher education. Their academic writing experience ranged from basic essay writing to research proposal writing. From the survey participants, 20 students were selected for interviews. The interview participants represented frequent, moderate, and limited users of artificial intelligence feedback.

The artificial intelligence tools used by students included grammar checkers, AI-powered writing assistants, paraphrasing tools, and generative AI chatbots. Students used these tools mainly to check

grammar, revise sentence clarity, improve vocabulary, receive suggestions on paragraph organization, paraphrase sentences, and obtain comments on overall writing quality. The study did not require students to use one specific tool because it aimed to examine authentic student practices with tools they commonly used in academic writing.

The quantitative instrument was the Artificial Intelligence-Assisted Feedback Experience Questionnaire. The questionnaire consisted of 36 items grouped into four dimensions: perceived usefulness, revision behavior, feedback trust, and perceived challenges. Each item used a five-point Likert scale ranging from 1, strongly disagree, to 5, strongly agree. Perceived usefulness measured students' views on the benefits of artificial intelligence feedback. Revision behavior measured how students used feedback in revising drafts. Feedback trust measured the extent to which students trusted and evaluated artificial intelligence suggestions. Perceived challenges measured difficulties related to accuracy, dependency, ethics, and authorial voice. The reliability coefficient of the questionnaire was 0.90, indicating strong internal consistency.

Qualitative data were collected through semi-structured interviews and document analysis. The interviews explored students' reasons for using artificial intelligence feedback, types of feedback received, revision decisions, perceived benefits, challenges, and ethical concerns. Each interview lasted approximately 30 to 45 minutes. Writing document analysis involved 48 student drafts consisting of initial drafts, artificial intelligence feedback notes, and revised drafts. The draft analysis examined feedback uptake patterns, including accepted feedback, modified feedback, rejected feedback, and uncritically adopted feedback.

Quantitative data were analyzed using descriptive statistics, including means, standard deviations, and percentages. The interpretation used the following categories: 1.00-1.80 very low, 1.81-2.60 low, 2.61-3.40 moderate, 3.41-4.20 high, and 4.21-5.00 very high. Qualitative data were analyzed thematically through data familiarization, coding, theme development, theme review, and interpretation. Writing drafts were analyzed by comparing original sentences, artificial intelligence suggestions, and revised versions. Data triangulation was conducted by comparing survey results, interview findings, and draft analysis. Ethical procedures were applied by obtaining informed consent, anonymizing student identities, and ensuring that students' participation did not affect course grades.

Several procedures protected the quality of the data. The questionnaire was reviewed by three experts in English language education and educational technology, with Aiken validity indices

between 0.81 and 0.93, and was piloted with 30 students outside the sample, where corrected item-total correlations ranged from 0.38 to 0.74 and the overall reliability reached a Cronbach's alpha of 0.89. Reliability for the four subscales was also acceptable: perceived usefulness 0.86, feedback trust 0.83, perceived challenges 0.85, and self-reported uptake 0.81. For the draft analysis, the four uptake categories were defined in a coding manual with authentic examples, ten drafts were coded jointly for calibration, and 30% of the corpus was then coded independently by two researchers, producing a Cohen's kappa of 0.84. Interview transcripts were returned to 12 participants for member checking, and questionnaire, draft, and interview evidence were compared through joint display analysis so that convergent and divergent findings could be identified explicitly.

C. Finding and Discussion

1. Finding

The findings are presented in four parts. The first part reports students' general experiences with artificial intelligence-assisted feedback. The second part explains students' perceived benefits. The third part presents challenges encountered by students. The fourth part reports students' feedback uptake patterns in academic writing revisions.

Students' General Experiences with Artificial Intelligence-Assisted Feedback

The survey findings showed that students had generally positive experiences with artificial intelligence-assisted feedback. The overall mean score was 3.91 with a standard deviation of 0.54, indicating a high level of perceived usefulness and engagement. Students reported that artificial intelligence tools helped them revise academic writing more quickly and gave them more confidence before submitting drafts.

Table 1 Students' General Experiences with Artificial Intelligence-Assisted Feedback

Dimension	Mean	SD	Interpretation
Perceived usefulness	4.18	0.51	High
Revision behavior	3.96	0.57	High
Feedback trust	3.42	0.66	High
Perceived challenges	4.05	0.59	High

Source: Processed from student questionnaire data.

The highest mean score appeared in perceived usefulness. Students considered artificial intelligence feedback helpful because it provided immediate responses and allowed them to revise drafts independently. However, the relatively lower score in feedback trust indicates that students did not fully trust every suggestion. Interview data showed that students often accepted grammar and clarity suggestions but hesitated when the tool suggested major changes in argument structure or meaning.

Perceived Benefits of Artificial Intelligence-Assisted Feedback

Students identified several benefits of artificial intelligence-assisted feedback. The most frequently reported benefits were grammar correction, sentence clarity, vocabulary improvement, revision confidence, and time efficiency.

Table 2 Perceived Benefits of Artificial Intelligence-Assisted Feedback

Benefit	Mean	SD	Interpretation
Improving grammar accuracy	4.35	0.49	Very high
Improving sentence clarity	4.27	0.53	Very high
Improving vocabulary choice	4.12	0.58	High
Supporting paragraph organization	3.82	0.64	High
Increasing revision confidence	4.10	0.55	High
Saving revision time	4.24	0.52	Very high

Source: Processed from student questionnaire data.

Grammar accuracy was the most valued benefit. Students used artificial intelligence tools to detect errors in verb tense, subject-verb agreement, articles, prepositions, punctuation, and sentence fragments. Many students stated that they felt more confident after checking their writing because the tool identified errors that they had missed.

Vocabulary improvement was another benefit. Students used artificial intelligence suggestions to replace repetitive words, choose more formal academic expressions, and avoid informal language. However, some students also reported that suggested vocabulary was sometimes too advanced or did not match the intended meaning.

Students perceived artificial intelligence feedback as less effective for paragraph organization than for grammar and sentence clarity. Although some tools offered suggestions about coherence, topic sentences, and transitions, students found these comments more general. They often needed lecturer feedback to understand whether their argument development was sufficient.

Challenges in Using Artificial Intelligence-Assisted Feedback

Despite positive experiences, students reported several challenges. These included inaccurate suggestions, generic feedback, difficulty evaluating feedback quality, overdependence, loss of authorial voice, and ethical uncertainty.

Table 3 Challenges in Using Artificial Intelligence-Assisted Feedback

Challenge	Mean	SD	Interpretation
Inaccurate or unsuitable suggestions	4.08	0.61	High
Generic feedback	3.96	0.67	High
Difficulty evaluating feedback quality	4.02	0.63	High
Overdependence on AI tools	4.11	0.56	High
Loss of authorial voice	3.88	0.69	High
Ethical uncertainty	4.23	0.54	Very high

Source: Processed from student questionnaire data.

Ethical uncertainty received the highest score among challenge items. Students were unsure about the acceptable boundary between using artificial intelligence for feedback and using artificial intelligence to rewrite or generate content. Some students believed grammar correction was acceptable but felt uncertain about paraphrasing, restructuring, or asking the tool to rewrite entire paragraphs.

Inaccurate suggestions also created difficulty. Students reported that artificial intelligence tools sometimes changed meaning, suggested unnatural academic phrases, or misunderstood discipline-specific terms. Some feedback improved grammatical form but weakened conceptual accuracy. This problem was especially visible in research writing, where terms and claims needed precision.

Overdependence emerged as a major concern. Some students admitted that they became less confident writing without artificial intelligence assistance. They used the tool before thinking deeply about revision. This indicates that artificial intelligence feedback may support writing performance but not necessarily writing development if students do not reflect on the feedback.

Loss of authorial voice was also reported. Students felt that heavily revised sentences sometimes sounded less like their own writing. Some revised drafts became more fluent but more generic. Students worried that their writing might no longer represent their own thinking.

Feedback Uptake Patterns in Student Drafts

The analysis of 48 writing drafts identified four feedback uptake patterns: direct acceptance, modified acceptance, selective rejection,

and uncritical adoption. Surface-level feedback was accepted more frequently than higher-order feedback.

Table 4 Feedback Uptake Patterns in Student Writing Drafts

Feedback Uptake Pattern	Frequency	Percentage	Dominant Feedback Type
Direct acceptance	312	45.75%	Grammar, spelling, punctuation
Modified acceptance	178	26.10%	Vocabulary and sentence clarity
Selective rejection	121	17.74%	Meaning, argument, technical terms
Uncritical adoption	71	10.41%	Rewritten sentences and paraphrases
Total	682	100%	

Source: Processed from writing draft analysis.

Modified acceptance occurred when students used artificial intelligence suggestions but changed parts of them to match their intended meaning. This pattern indicated stronger feedback literacy because students did not treat the suggestion as final. They adapted feedback based on context.

Selective rejection occurred when students decided not to use feedback because it changed meaning, sounded unnatural, or did not match the assignment. This pattern was more common among students with stronger writing competence. These students treated artificial intelligence feedback as input rather than authority.

Uncritical adoption occurred when students accepted rewritten sentences or paragraphs without understanding the changes. This pattern was risky because some revisions changed argument meaning, removed specific details, or produced overly generic academic prose. It was more common among students who lacked confidence in their own writing.

Qualitative analysis of interviews and drafts produced five major themes: immediate revision support, private error awareness, selective trust, ethical hesitation, and need for teacher guidance.

Table 5 Qualitative Themes on Students' Experiences

Theme	Evidence from Student Experience	Pedagogical Meaning
Immediate revision support	Students received feedback instantly before submission	AI feedback supported timely revision
Private error awareness	Students checked errors without public embarrassment	AI feedback reduced writing anxiety

Selective trust	Students trusted grammar feedback more than content feedback	Feedback literacy shaped uptake
Ethical hesitation	Students were unsure about acceptable AI use	Clear academic policy was needed
Need for teacher guidance	Students still needed lecturers for argument and source use	Human feedback remained essential

Source: Synthesized from interviews and draft analysis.

2. Discussion

The findings show that artificial intelligence-assisted feedback was perceived positively by students, especially for grammar correction, sentence clarity, vocabulary improvement, and time efficiency. This supports previous research indicating that automated writing evaluation and AI-powered writing assistants can provide immediate and individualized feedback for writing development (Warschauer & Ware, 2006; Ranalli, 2018; Nazari, Shabbir, & Setiawan, 2021). Students valued artificial intelligence feedback because it helped them revise drafts quickly and independently.

However, the findings also show that students' experiences were not uniformly positive. Artificial intelligence feedback was most useful for surface-level revision but less reliable for higher-order writing concerns. Grammar, spelling, punctuation, and clarity suggestions were frequently accepted, while feedback on argument structure, evidence, source integration, and disciplinary meaning required more caution. This finding aligns with Lee's (2020) argument that automated writing evaluation has utility but also limitations in EFL writing classrooms.

The difference between surface-level and higher-order feedback is important. Academic writing is not only about linguistic accuracy. Students need to construct arguments, synthesize literature, evaluate evidence, and maintain disciplinary meaning. Artificial intelligence tools may improve formal fluency but cannot fully understand the student's research intention, disciplinary context, or assignment expectations. Therefore, artificial intelligence feedback should complement, not replace, lecturer feedback.

Students' feedback trust was moderate compared with perceived usefulness. This indicates that students recognized both benefits and limitations. Some students used artificial intelligence feedback critically, accepting suggestions only when they matched their intended meaning. This behavior reflects feedback literacy. Bai and Hu (2017) similarly found that students could respond selectively to fallible automated feedback when they were aware of its strengths and

weaknesses. In this study, students with stronger writing competence showed more selective uptake.

Ethical uncertainty emerged as the strongest challenge. Students were unsure where acceptable assistance ended and academic misconduct began. This finding is consistent with recent discussions on generative artificial intelligence in education, which highlight issues of academic integrity, transparency, authorship, and data privacy (Barrot, 2023; Kasneci et al., 2023; Nguyen et al., 2023). Students need clear institutional guidelines about what forms of AI assistance are allowed, how to disclose AI use, and what types of tasks should remain independently produced.

The analysis of feedback uptake shows that direct acceptance was common, particularly for grammar and mechanics. This is not necessarily problematic if feedback is accurate and students understand the correction. However, uncritical adoption of rewritten sentences is more problematic because students may submit text they cannot explain. This affects both academic integrity and learning. Students should be encouraged to ask why a suggestion improves the text, not merely whether it sounds better.

Comparison with Previous Studies

The pattern reported here is consistent with the automated-feedback literature. Warschauer and Ware (2006) argued early that automated writing evaluation should be judged by what learners do with the feedback rather than by the accuracy of the engine, and later studies have borne this out. Ranalli (2018) found that students could act on automated corrective feedback but often did so without understanding the underlying rule, and Koltovskaia (2020) showed that behavioural engagement with automated feedback was high while cognitive engagement remained shallow. Lee (2020) reached a similar conclusion for EFL classrooms, noting that automated evaluation is useful for form but weak for meaning. Bai and Hu (2017) add an important qualification: when the feedback is wrong, stronger writers reject it and weaker writers accept it, which is exactly the asymmetry the present study observed between selective rejection and uncritical adoption.

The advantage the present participants reported for surface-level feedback also matches the written-corrective-feedback tradition. Ferris (1999, 2004) argued that targeted correction helps learners who know what to look for, and Bitchener and Knoch (2010) and Shintani, Ellis, and Suzuki (2014) demonstrated measurable accuracy gains for specific structures under controlled conditions. Hyland and Hyland (2006) had already noted, however, that feedback works through the

relationship between the reader and the writer, which is precisely what an automated reader cannot supply. Zhang and Hyland (2018, 2022) show that students engage more deeply with teacher feedback than with automated feedback even when the two overlap in content, and Dawson et al. (2019) report that students and teachers judge feedback quality by different criteria. The moderate feedback-trust score obtained in this study should be read against that background.

Studies of generative tools report the same double-sided picture. Nazari, Shabbir, and Setiawan (2021) found in a randomised controlled trial that an AI writing assistant improved writing performance and self-efficacy, and Wei (2023) reported comparable gains for Chinese EFL learners. Godwin-Jones (2022) and Strobl et al. (2019) show that such tools are best understood as one layer in a wider ecology of digital writing support rather than as a replacement for instruction. On the generative side, Barrot (2023) and Kohnke, Moorhouse, and Zou (2023) set out both the potential and the pitfalls of ChatGPT for second language writing, while Chan and Hu (2023) found that students value the speed of generative feedback but worry about accuracy and academic integrity, echoing the ethical uncertainty that scored highest among the challenges in this study. Nguyen, Hong, Dang, and Huang (2024) trace the actual sequences of human-AI interaction during academic writing and show that productive collaboration involves iterative negotiation rather than wholesale acceptance, which corresponds to the modified-acceptance pattern identified here.

Theoretical and Pedagogical Implications

Theoretically, the findings support a feedback-literacy reading of artificial intelligence in academic writing. What distinguished the students who benefited from those who did not was not access to the tool, which was identical, but the capacity to evaluate a suggestion before acting on it. The four uptake patterns identified here can be arranged along that dimension: selective rejection and modified acceptance both require a judgement about whether the suggestion serves the writer's meaning, while direct acceptance and uncritical adoption suspend that judgement. This also explains why the tool helped most at the level of grammar and sentence clarity and least at the level of argument and organisation, since only the former can be evaluated without reference to the writer's overall rhetorical purpose.

Pedagogically, three implications follow. First, artificial intelligence-assisted feedback should be introduced together with explicit instruction in how to interrogate a suggestion, for example by requiring students to record why a suggestion was accepted, modified,

or rejected. Second, the division of labour between the tool and the teacher should be made explicit: the tool is efficient for surface-level revision, while teacher and peer feedback remain necessary for argument, coherence, and disciplinary conventions. Third, institutions need a clear and specific policy on acceptable use. Ethical uncertainty was the highest-scoring challenge in this study, and it arose from the absence of guidance rather than from any intention to cheat; a policy that distinguishes language editing from content generation would resolve most of it.

Limitations and Directions for Future Research

Several limitations qualify these findings. First, the study involved students from one faculty at one university, so the patterns described here should be replicated before wider claims are made. Second, the questionnaire measured perceptions rather than performance, and perceived usefulness is not a proxy for writing gain; the draft analysis partly compensates for this but covers a single assignment cycle. Third, students used more than one tool and the study did not control which, so differences between systems could not be separated. Fourth, the design was cross-sectional and cannot show how uptake patterns change as students become more experienced. Future research could compare specific tools under controlled conditions, track the same writers across several assignments, test whether explicit instruction in feedback literacy shifts uptake from direct to modified acceptance, and examine how institutional policy on generative tools affects what students actually do.

D. Conclusion

This study examined students' experiences and challenges in using artificial intelligence-assisted feedback in academic writing. The findings showed that students perceived artificial intelligence feedback as useful for grammar correction, sentence clarity, vocabulary improvement, revision confidence, and time efficiency. However, students also experienced challenges related to inaccurate suggestions, generic comments, difficulty evaluating feedback quality, overdependence, loss of authorial voice, and ethical uncertainty. Writing draft analysis showed that students accepted surface-level feedback more frequently than higher-order feedback. Direct acceptance was most common for grammar and mechanics, while selective rejection occurred when feedback changed meaning or weakened argument accuracy. The study concludes that artificial intelligence-assisted feedback can support academic writing when used critically, selectively, and transparently. It should be treated as a

revision aid rather than an authority or substitute writer. Students need AI feedback literacy, including the ability to evaluate suggestions, preserve authorial voice, and disclose AI assistance responsibly. Lecturers and institutions should provide clear policies, guided practice, and writing pedagogy that integrates artificial intelligence without reducing students' responsibility for learning. Future research should examine longitudinal development, compare different AI feedback tools, and investigate how AI-assisted feedback affects students' academic writing quality across disciplines.

Declaration of AI-Assisted Technologies

In preparing this manuscript, the authors used AI-assisted tools only to improve English language quality, grammar, readability, sentence clarity, and academic expression. The AI tools were not used to generate research data, conduct data analysis, produce findings, or draw conclusions. All conceptual arguments, data interpretation, methodological decisions, and scholarly conclusions remain the full responsibility of the authors. The authors carefully reviewed, edited, and verified all AI-assisted outputs to ensure the accuracy, integrity, and originality of the final manuscript.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Bibliography

- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67-81. doi: 10.1080/01443410.2016.1223275
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. doi: 10.1016/j.asw.2023.100745
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207-217. doi: 10.1016/j.jslw.2010.10.002
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. doi: 10.1080/02602938.2018.1463354
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. doi: 10.1186/s41239-023-00411-8
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research

- recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. doi: 10.1016/j.caeai.2022.100118
- Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2019). What makes for effective feedback: Staff and student perspectives. *Assessment & Evaluation in Higher Education*, 44(1), 25-36. doi: 10.1080/02602938.2018.1467877
- Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott. *Journal of Second Language Writing*, 8(1), 1-11. doi: 10.1016/S1060-3743(99)80110-6
- Ferris, D. R. (2004). The grammar correction debate in L2 writing: Where are we, and where do we go from here? *Journal of Second Language Writing*, 13(1), 49-62. doi: 10.1016/j.jslw.2004.04.005
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5-24.
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101. doi: 10.1017/S0261444806003399
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. doi: 10.1016/j.lindif.2023.102274
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELJ Journal*, 54(2), 537-550. doi: 10.1177/00336882231162868
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback in ESL writing. *Assessing Writing*, 44, 100450. doi: 10.1016/j.asw.2020.100450
- Lee, I. (2020). Utility and limitation of automated writing evaluation in EFL writing classrooms. *Language Teaching Research*, 24(6), 777-798. doi: 10.1177/1362168819826241
- Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centered framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45(4), 527-540. doi: 10.1080/02602938.2019.1667955
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), e07014. doi: 10.1016/j.heliyon.2021.e07014
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847-864. doi: 10.1080/03075079.2024.2323593
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28, 4221-4241. doi: 10.1007/s10639-022-11316-w

- Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning*, 31(7), 653-674. doi: 10.1080/09588221.2018.1428994
- Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103-131. doi: 10.1111/lang.12029
- Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33-48. doi: 10.1016/j.compedu.2018.12.005
- Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, 10(2), 157-180. doi: 10.1191/1362168806lr190oa
- Wei, P. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Frontiers in Psychology*, 14, 1249991. doi: 10.3389/fpsyg.2023.1249991
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39. doi: 10.1186/s41239-019-0171-0
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90-102. doi: 10.1016/j.asw.2018.02.004
- Zhang, Z. V., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51, 100586. doi: 10.1016/j.asw.2021.100586