TESL-EJ Teaching English as a Second or Foreign Language
The Electronic Journal for English as a Second Language

AI vs. Human Feedback in EFL Writing: Examining the Interplay Between Feedback Modalities and Digital Proficiency

*****On the Internet*****

May 2026 – Volume 30, Number 1
https://doi.org/10.55593/ej.30117int2

Veni Roza[1]
<veniroza@uinbukittinggi.ac.id>

Melyann Melani[1]
<melyannmelani@uinbukittinggi.ac.id>

Reflinda[1]
<reflinda@uinbukittinggi.ac.id>

Rianto[1]
<rian.anto002@gmail.com>

Ridha Ahida[1]
<ridhaahida@uinbukittinggi.ac.id>

Safnil Arsyad[2]
<safnil@unib.ac.id>

Genta Sakti[1]
<genta.sakti@uinbukittinggi.ac.id>

Yuyud Andam Mellian[1]
<yuyudmellian45@gmail.com>

[1] UIN Sjech M. Djamil Djambek Bukittinggi, Indonesia
[2] Bengkulu University, Indonesia

Abstract

The study explores the differential effects of AI-generated, teacher, and peer feedback on improving students' writing development with digital proficiency as the factorial variable. A quasi-experimental mixed-methods approach was administered to 150 Indonesian university undergraduate English education students divided into high, moderate, and low digital proficiency levels. Pre- and post-tests, a digital proficiency survey, and semi-structured interviews were used to collect data. Statistical analyses, e.g., paired-sample t-tests, Generalized Linear Mixed Model (GLMM), and independent t-tests, established that teacher feedback generated the largest improvements on content (d = 0.41, p = 0.006) and grammar (d

= 0.45, p = 0.002) but also that AI-distributed feedback had its most significant effect on the grammar (d = 0.34, p = 0.025). Peer feedback helped in content generation (d = 0.36, p = 0.013) but had little influence on structural accuracy. GLMM results indicated a significant interaction between feedback type and digital proficiency (F(2, 143) = 3.410, p = .036), demonstrating that the effectiveness of the feedback modality was moderated by the student's technological readiness. Specifically, students with higher digital proficiency scores showed a more significant positive response to AI-generated feedback, leveraging its immediacy for structural revisions. Qualitative findings indicated that high digital proficiency learners valued AI feedback for being quick and transparent while low digital proficiency students found teacher feedback more structured and comprehensible. The findings identify the importance of including diversified types of feedback in line with students' digital proficiency for writing improvement in EFL contexts.

Keywords: AI feedback; digital proficiency; peer feedback; teacher feedback.

Research in applied linguistics and second language writing has consistently demonstrated the crucial role of feedback in improving learners' writing performance, particularly in areas such as accuracy, coherence, and content development (Hyland & Hyland, 2006; Liu & Yu, 2022). Studies comparing feedback types accentuate the strengths and limitations of each approach: teacher feedback is often valued for its clarity and authority (Teng & Ma, 2024; Wang & Han, 2022); peer feedback promotes collaboration and critical engagement (Park, 2025; Zou et al., 2023); and automated feedback, especially AI-based tools such as ChatGPT, shows growing potential in delivering immediate, structured responses to student writing (Steiss et al., 2024; Teng, 2024a, 2024b, 2025c). Although integrated feedback methods appear to enhance revision quality and learner engagement (Tian & Zhou, 2020; Zhang & Hyland, 2022), little is known about how these feedback types perform across specific writing dimensions such as content, organization, grammar, vocabulary, and mechanics, especially when used concurrently.

In the Indonesian EFL context, writing development faces persistent challenges. Students often struggle with limited vocabulary, weak organization, lack of critical thinking, and frequent grammatical errors, all exacerbated by a test-oriented educational culture and limited exposure to authentic materials. Large class sizes and teachers' heavy workloads further restrict the quality and frequency of feedback (Masrul et al., 2024). Moreover, cultural hesitation toward peer critique undermines the effectiveness of collaborative feedback. Emerging technologies such as ChatGPT offer an opportunity to supplement traditional approaches by providing instant, personalized feedback (Arifin et al., 2024; Oktarina et al., 2024). Nonetheless, the benefits of AI-based tools are unevenly distributed due to the digital divide, particularly between urban and rural students, which limits access and the ability to fully utilize these tools (Achruh et al., 2024; Mustopa et al., 2024).

Despite increasing attention to AI-generated feedback, existing research rarely explores how digital proficiency influences learners' engagement with various feedback types. Although AI tools can offer accurate grammatical and structural corrections (Alghannam, 2024; Escalante et al., 2023; Rahmi et al., 2024), concerns remain about their ability to support deeper content development and critical thinking (Evmenova et al., 2024). Furthermore, learners' attitudes toward technology and levels of digital literacy can mediate their feedback experiences and learning outcomes (Alamäki et al., 2024; Saúde et al., 2024). In light of these gaps, this study examines the combined and comparative effects of teacher, peer, and AI-generated feedback on students' writing achievement and explores how digital proficiency and proficiency levels interact with feedback effectiveness. By focusing on these dynamics in the Indonesian EFL context, the study aims to contribute to more inclusive, context-sensitive feedback practices

that bridge technological and pedagogical divides. The following research questions guide the study:

1. How do different feedback types (GPT, teacher, peer) influence students' writing improvements across various criteria (content, organization, grammar, vocabulary, mechanics)?
2. How do feedback types and students' digital proficiency interact to impact the effectiveness of feedback on their writing outcomes?
3. How do students' digital proficiency shape their preferences for, perceptions of, and ability to apply different types of feedback (GPT, teacher, peer)?

Literature Review

The Role of Feedback in EFL/ESL Writing Development

Feedback plays a central role in the development of EFL/ESL writing by supporting learners' ability to revise iteratively, refine linguistic accuracy, and enhance vocabulary use. It not only helps identify and correct errors but also fosters rhetorical development and internalization of language forms (Liu & Yu, 2022). Teacher-provided contextual feedback has been shown to significantly improve accuracy and rhetorical competence (Li & Zhang, 2022), while dynamic, dialogic feedback facilitates long-term linguistic gains (Apridayani & Waluyo, 2025). Nonetheless, feedback is not universally effective; its impact is shaped by variables such as learner characteristics, classroom context, and instructional methods (Evans et al., 2010; Hyland & Hyland, 2006; Wahyudi & Sari, 2016), making it a socially embedded and context-sensitive pedagogical practice.

Theoretical perspectives identify feedback's dual function in language development. From a sociocultural standpoint, feedback acts as scaffolded support that enables learners to internalize linguistic structures (Poehner & Leontjev, 2020; Villamil & de Guerrero, 2006), whereas cognitive theories, e.g., Krashen's Input Hypothesis, stress its role in promoting noticing and language processing (Leow & Driver, 2021; Kartchava, 2013; Nassaji, 2021). Lyster (2012) further emphasizes feedback's multifaceted nature, involving cognitive, social, and pedagogically contested, especially regarding mode, timing, and cognitive load. In contemporary classrooms, feedback delivery now includes not only teacher input but also peer evaluation and AI-generated suggestions, each offering unique affordances for supporting diverse learner needs and instructional goals.

Teacher Feedback: Effectiveness and Challenges

Teacher-written feedback (TWF) remains a central strategy for enhancing EFL/ESL students' writing proficiency, particularly in higher education contexts. Studies from Thailand and China confirm that TWF provides reliable guidance for linguistic and rhetorical improvement (Cheng & Liu, 2022; Waluyo & Apridayani, 2024). However, challenges persist, i.e., grammar-focused feedback demands considerable time and may conflict with institutional expectations or pedagogical beliefs (Lee, 2019; Bailey & Garner, 2010). Student uptake of feedback varies based on their proficiency and the type of feedback received, prompting calls for tailored and student-centered approaches (Cheng & Liu, 2022). Researchers have emphasized the importance of directed feedback strategies, integrating peer feedback, and managing learner expectations to enhance its impact (Lee, 2019). A sociocultural view further advocates for process-based feedback that fosters learner autonomy and reflective engagement (Lee, 2014). Nonetheless, disparities in perceptions and practices between teachers and students remain underexplored, particularly in diverse higher education contexts (Mamad & Vígh, 2024).

In Indonesia's resource-constrained educational settings, the effectiveness of TWF is often limited by large class sizes and insufficient institutional support. These factors hinder the provision of individualized and consistent feedback, weakening its formative potential. Despite these constraints, Indonesian learners express positive attitudes toward TWF across both online and face-to-face modalities, valuing its role in guiding revision and improving writing outcomes (Linuwih et al., 2024; Mafulah & Cahyono, 2023). The systemic limitations of TWF emphasize the need for supplementary feedback mechanisms, such as peer assessment and AI-driven tools, to support equitable and scalable writing instruction. Addressing these challenges is crucial for fostering an inclusive feedback culture that supports diverse learner needs while maintaining instructional quality.

Peer Feedback: Benefits and Limitations

Peer feedback, grounded in collaboration theories and constructivist learning, offers a student-centered approach to writing development through mutual interaction (Palincsar, 1998). It has been shown to improve motivation, critical thinking, and revision engagement, particularly in EFL contexts, e.g., Saudi Arabia and Palestine (Alharbi, 2021; Farrah, 2012). Technology-mediated peer feedback can further reduce anxiety, making it more adaptable to diverse learning environments (Peungcharoenkun & Waluyo, 2024). However, students' ability to give or accept peer feedback is influenced by proficiency levels, confidence, and cultural factors such as face-saving norms, which are particularly relevant in Indonesian classrooms (Yu & Lee, 2016; Damanik, 2022). While teacher guidance can improve feedback quality, challenges persist due to inconsistent feedback and limited training in peer review strategies (Suherman, 2018). Despite these limitations, peer feedback remains valuable when supported by adequate instructional scaffolding.

In the Indonesian context, peer review in English writing classes presents mixed outcomes. Group-based peer activities often yield minimal impact on writing improvement without structured support (Kusumaningrum et al., 2019). Although students tend to trust teacher feedback more, peer feedback can still foster writing development if properly implemented (Damanik, 2022). Effective peer review requires teacher-led training to equip students with the skills and mindset needed to provide meaningful evaluations (Damanik, 2022). Given the variability in peer feedback quality, there is growing interest in examining how AI-generated feedback compares in supporting writing improvement. Future research should explore the intersection of peer and AI-based feedback systems to determine how each can serve as a complementary tool tailored to learners' diverse needs in digital learning environments.

AI-Generated Feedback: Potential and Limitations

AI-generated feedback, particularly through tools such as ChatGPT, has emerged as a promising supplement to traditional feedback in EFL writing instruction. Studies conducted in Hong Kong and Macau show that AI feedback can improve student engagement and writing performance (Chan et al., 2024; Teng, 2024b). Indonesian research echoes these findings, with evidence that AI tools enhance grammar and sentence structure, leading to improved writing outcomes (Marmoah et al., 2024; Rafida et al., 2024). Students appreciate the immediacy and structure of AI responses, which contribute to greater writing confidence and efficiency (Teng, 2024b; Gozali et al., 2024). However, while AI feedback builds feedback literacy and positively shifts learner attitudes, concerns remain regarding creativity suppression, potential overreliance, and academic integrity (Rafida et al., 2024; Evenddy, 2024). Integrating AI within human-supervised settings is recommended to balance innovation with pedagogical control (Escalante et al., 2023).

Despite its promise, AI-generated feedback presents notable limitations. Studies draw attention to issues with accuracy, lack of personalization, and inconsistencies in AI responses, especially among linguistically diverse learners (Alghannam, 2024; Evmenova et al., 2024; Huang & Teng, 2025). Comparative studies offer mixed findings; for instance, some report no significant differences between AI and human feedback outcomes (Escalante et al., 2023), although others emphasize the superior contextual richness of expert feedback (Jansen et al., 2024). Furthermore, while AI offers structured, descriptive suggestions, peer feedback often better identifies salient writing issues (Banihashem et al., 2024; Waluyo & Kusumastuti, 2024). Given these complexities, a hybrid model that combines AI-generated, teacher, and peer feedback may offer the most effective approach, supporting both immediate learner needs and long-term writing development in contexts such as Indonesia.

AI, Teacher, and Peer Feedback in EFL Writing: Evidence, Limits, and Design Implications

Comparative scholarship increasingly portrays AI-, teacher-, and peer-mediated feedback as complementary rather than competing resources in EFL writing. Reviews suggest that AI-supported learning offers immediacy and scalable, criterion-referenced comments, yet it remains uneven in accuracy and limited in dialogic, communicative, and affective support; as a result, human monitoring and pedagogical orchestration are essential (Alghannam, 2024; Yang & Kyun, 2022). Technology-enhanced teacher and peer feedback generally outperform self-feedback in collaborative writing and, alongside self- and peer review, cultivate critical thinking and engagement (Zou et al., 2023). At the same time, implementation challenges persist, from platform design and classroom integration to the still nascent understanding of AI's effects on learners' emotions and motivation, accentuating the need for robust, context-sensitive designs and further empirical work on affective outcomes (Alshumaimeri & Alshememry, 2023; AlTwijri & Alghizzi, 2024). Taken together, current evidence favors a hybrid model in which AI augments, but does not replace, structured teacher guidance and peer interaction (Yang & Kyun, 2022).

Recent empirical studies deepen this picture by showing how outcomes hinge on task design, training, and learners' evolving engagement with feedback sources. Well-supported platforms and explicit training enable technology-mediated peer feedback to enhance motivation and attitudes toward writing, while learners' uptake varies dynamically with individual and contextual factors (Cuocci et al., 2023; Tian & Zhou, 2020). Although students may perceive AI feedback as more "expert" or trustworthy than input from educators or peers, this perception does not obviate the distinctive value of human feedback for higher-order reasoning, audience awareness, and socio-emotional support (Ruwe & Mayweg-Paus, 2023). Evidence from complex writing tasks indicates that tools such as ChatGPT can supply rapid, content-focused suggestions but function most productively alongside peer review and teacher mediation, not as a standalone solution (Alghannam, 2024; Banihashem et al., 2024; Teng & Huang, 2025; Yang & Kyun, 2022). Accordingly, persuasive EFL feedback ecosystems sequence AI for preliminary diagnosis and practice, harness peer dialogue for meaning negotiation and rhetorical refinement, and rely on teacher intervention to calibrate quality, ensure ethical use, and cultivate reflective, transferable writing strategies.

Digital proficiency and Their Impact on Feedback Effectiveness

Students' digital proficiency is a decisive precondition for realizing the benefits of AI-mediated feedback in EFL writing. Learners who hold positive attitudes toward digital tools and possess stronger operational skills are more likely to interpret, prioritize, and implement AI suggestions effectively; by contrast, skepticism and limited skills constrain uptake and can amplify concerns about accuracy, ethics, and the impersonality of machine responses (Getenet et al., 2024; Shen et al., 2023; Solak, 2024; Zhang & Hyland, 2023). While AI tools can streamline

drafting and revision and increase learners' willingness to engage with feedback, many students, especially those with lower proficiency or weaker digital literacy, still require human mediation to contextualize comments, connect them to task goals, and calibrate the level of support to their needs (Pang et al., 2024; Rad et al., 2024). Accordingly, a hybrid feedback ecology, in which teachers explicitly teach "feedback literacy," provide targeted scaffolds for interpreting AI output, and decide when to prioritize human over automated input, offers a more equitable and inclusive pathway for diverse learners (Chan & Hu, 2023).

At the design level, evidence on digital feedback modalities converges on two principles: match the tool to the task, and stage feedback to align with the writing process. Meta-analytic work on computer-mediated feedback indicates large overall effects, with automated and immediate feedback exerting the strongest influence, likely because rapid, criterion-specific input increases revision frequency and reduces ambiguity at early drafting stages (Park, 2025). Automated written corrective feedback (AWCF) and computer-mediated peer feedback (CMPF) reliably improve text quality and foster autonomy by distributing evaluative authority and encouraging self-regulation, but their effectiveness is conditioned by learner proficiency, genre demands, cognitive load, and timing across the drafting cycle (Allen & Mills, 2016; Hojeij & Ayber, 2022; Park, 2025; Tan et al., 2023). Electronic peer feedback can strengthen sentence complexity and grammatical accuracy, yet its benefits are uneven without clear rubrics, training, and norms for constructive dialogue (Shang, 2022). The most consistent gains arise when AWCF is sequenced for quick, surface-level diagnostics and CMPF is used for dialogic elaboration of ideas, with teachers orchestrating the sequence, modeling how to triage machine comments, and scheduling intervention points when substantive revision is still feasible (Tan et al., 2023).

Methods

Research Design, Context, Intervention, and Procedures

This mixed-methods study examined the comparative effectiveness of ChatGPT, teacher, and peer feedback on the writing performance of Indonesian EFL university students, while also exploring how digital proficiency moderated these effects. Conducted at Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi with 150 English Education undergraduates (117 females, 33 males), the study responded to persistent challenges in Indonesia's EFL context, including large class sizes, limited feedback access, and digital inequality (Masrul et al., 2024; Mustopa et al., 2024). Participants were categorized into high (n = 30), moderate (n = 97), and low (n = 23) digital proficiency groups using a pre-intervention literacy survey, allowing for analysis of how digital skills shaped students' engagement with feedback across modalities. Participants were recruited from three intact course sections via instructor announcements; participation was voluntary with informed consent. In addition to gender, we recorded age and self-reported prior academic-writing coursework/experience via a brief intake form to contextualize the sample; these descriptors were used descriptively.

A quasi-experimental design was employed to examine causal relationships in an authentic classroom environment where random assignment was impractical (Gopalan et al., 2020). Participants were assigned to three instructional groups: ChatGPT feedback (Class A), teacher feedback (Class B), and peer feedback (Class C). Because cohorts were pre-formed by the registrar, random assignment at the student level was not feasible; intact sections were therefore assigned at the class level to minimize cross-condition contamination. To mitigate confounding, we standardized the syllabus, writing prompts, time-on-task, submission timelines, and assessment rubrics across sections; provided common preparatory instruction; and scheduled tasks in parallel. We treat the study as a non-equivalent groups pretest–posttest design and interpret effects with appropriate caution. The intervention, conducted over 10

weeks, involved iterative argumentative writing tasks with structured pre- and post-tests assessing writing performance across five criteria: content, organization, grammar, vocabulary, and mechanics. Class A received AI-generated feedback via ChatGPT (GPT-4.0), which provided immediate, automated suggestions focusing on grammar, sentence clarity, and structure, which are the features recognized for their efficiency but limited depth in rhetorical feedback (Alghannam, 2024; Chan et al., 2024; Teng, 2024b).

AI feedback was operationalized with a standardized prompt aligned to the analytic rubric; students pasted their draft and rubric, and ChatGPT returned (a) strengths, (b) three prioritized, example-based suggestions, and (c) a short revision plan; the model was instructed not to rewrite the essay. Class B received personalized written feedback from experienced EFL instructors, targeting content development, coherence, and language accuracy, in line with best practices in teacher-mediated formative assessment (Lee, 2019; Waluyo & Apridayani, 2024). Teachers followed a written protocol adapted from Lee (2019) that prioritized global issues (content, organization) before local form, used selective marking, and required feedforward comments specifying concrete next steps. Class C engaged in structured peer review using guided rubrics to evaluate organization and language use, a method known to enhance critical thinking and promote learner autonomy (Farrah, 2012; Yu & Lee, 2016; Palincsar, 1998). Peers received brief training in "feedback literacy" and calibration using anchor essays; sentence stems and a comment template were provided to promote constructive, specific, and actionable feedback.

To support intervention trustworthiness and ensure instructional consistency, all participants received training on argumentative writing structure, digital tool navigation, and feedback literacy. Fidelity was monitored with a short adherence checklist and spot audits of feedback samples in each condition. In addition to quantitative measures, semi-structured interviews were conducted post-intervention to explore students' experiences, perceptions, and emotional-cognitive engagement with different feedback modalities. The integration of quantitative and qualitative findings allowed for a robust understanding of how each feedback type, within a digitally uneven educational context, affects writing development and student agency. The study responds to growing calls for feedback models that are pedagogically grounded, scalable, and inclusive of both technological and human dimensions in EFL instruction (Escalante et al., 2023; Zhang & Hyland, 2022).

Table 1. Comparative Characteristics of Feedback Across Groups

Group	Feedback Type	Focus Areas	Timing	Strengths	Limitations
Class A	AI-Generated ChatGPT 4.0	Grammar, sentence clarity, structure	Immediate (post-submission)	Fast, consistent, scalable; reduces teacher workload	Limited depth in content; lacks nuance in rhetorical and contextual feedback
Class B	Experienced EFL instructors	Content develop-ment, coherence, mechanics, vocabulary	24–48 hours after submission	Personalized, context-sensitive, addresses higher-order writing skills	Time-consuming; limited scalability due to workload and class size
Class C	Peer Feedback	Organization, grammar, mechanics	In-class peer review sessions	Encourages autonomy, collaboration, and feedback literacy	Quality varies based on peer training, confidence, and cultural norms

Research Instruments and Measures

Digital proficiency Survey. Students completed the digital proficiency survey, inspired by established models (Jesudass & Mirunalini, 2023), prior to the start of the intervention. The survey assessed six key dimensions (60 items), which included learning engagement with 11 items and lesson planning with 10 items, along with instructional material preparation, which consisted of 10 items, and teaching techniques application with 8 items, while communication and evaluation were assessed with 7 items and 14 items, respectively, using a scale from

strongly disagree to strongly agree. The survey results grouped participants into three categories based on their scores: high proficiency (above 191), moderate proficiency (between 170 and 190), and low proficiency (below 169), which helped analyze how digital proficiency affected their ability to receive and use feedback. Cut-scores were set a priori based on the score distribution and interpretive guidelines for the instrument.

Pre- and Post-Writing Assessments. To assess students' initial writing skills, all the participants finished a pre-writing test at the start of the study. Students wrote essays that were scored analytically for the content organization, grammar, vocabulary, and mechanics. The analytic scoring approach was preferred to holistic methods, as this method gives full feedback concerning students' areas of strength and areas for improvement in the various components of writing (Nodoushan, 2014). After the feedback session, the students did a post-writing test under the same conditions as those in the initial assessment. Prompts, time limits, and test conditions were identical across groups; the same analytic rubric used for instruction was applied at both time points.

Semi-Structured Interviews. The research group conducted semi-structured interviews with 15 participants to assure a balanced representation among different digital proficiency groups while complementing the quantitative findings. Researchers looked into students' preferences about the kinds of feedback provided for and about the effectiveness and ease of use of feedback during writing revisions. The team transcribed the qualitative data together with the coding and thematic analysis to surface the patterns of feedback engagement. Main themes identified in analyzing feedback included students' and teachers' perceptions of clarity of the feedback; the main bottlenecks about actually getting suggestions applied; and students' levels of trust in feedback from AI systems against those in their teachers and peers. The interview insights were further combined with the quantitative findings to provide a thorough picture about how feedback works for students with different digital proficiencies. Interviews followed a semi-structured protocol (e.g., perception, use, mediation, and trust), lasted approximately 20–30 minutes, and were conducted after the post-test to avoid priming effects.

Data Analysis

In analyzing the pre- and post-test scores, paired-sample t-tests were conducted to test writing before and after treatment within each feedback group, while the effect size (Cohen’s d) was calculated for practical significance of changes seen. GLMM was then conducted along with Welch’s t-tests to measure the interaction of the two factors-feedback type and digital proficiency-in influencing writing improvements in the students. Primary outcomes were gain scores (post–pre) computed for each analytic criterion and the total score. The GLMM evaluated the main effects of feedback modality and digital proficiency alongside their interaction term to identify how digital proficiency moderated feedback effectiveness.; Welch's t-tests were used for comparisons because the datasets were reasonably normal and Welch's t-test is a more conservative test that accounts for inequality in variance. Following this, the transcribed data from the semi-structured interview were coded for thematic analysis, with the identification of key themes associated with preferences for feedback, ease of understanding, and application of feedback.

Reliability and Validity Procedures

An excellent internal consistency score ($\alpha = 0.93$) confirmed the reliability of the Students’ Digital Proficiency Survey, which comprised 60 items. The subscales for Learning Engagement ($\alpha = 0.75$) and Evaluation ($\alpha = 0.74$) demonstrated good reliability, while Lesson Planning ($\alpha = 0.73$) and Instructional Material Preparation ($\alpha = 0.70$) were considered acceptable. Application of Teaching Techniques ($\alpha = 0.68$) and Communication ($\alpha = 0.67$) fell

within the lower acceptable range, and these limitations were considered when interpreting subscale-level findings. Content validity was established through expert review, and the use of a representative sample from the English Education Department supported broader generalizability. Besides, triangulation with qualitative data enhanced the overall validity of the study through cross-verification.

Results

Quantitative Findings

Analysis of Writing Performance Improvement Across Feedback Modalities. Table 2 presents a summary of the analysis of Writing Performance Improvement Across Feedback Modalities. The following paragraphs explain the improvements in writing criteria across feedback types and provide a comparative analysis of total and final score improvements.

Table 2a. Results of the students' pre- and post- writing tests across feedback types

Feedback Type	Criterion	Pre M	Pre SD	Post M	Post SD	*t*	*p*	Effect Size
GPT	**Content**	**2.56**	**1.05**	**2.88**	**0.90**	**−2.22**	**0.031**	**0.30**
	Organization	2.62	1.16	2.92	0.83	−1.82	0.075	0.26
	Grammar	**2.40**	**1.07**	**2.76**	**0.85**	**−2.31**	**0.025**	**0.34**
	Vocabulary	2.52	1.05	2.56	0.79	−0.25	0.802	0.04
	Mechanics	2.62	1.09	2.90	0.89	−1.61	0.114	0.26
TEACHER	Content	2.50	1.04	2.64	0.83	−1.07	0.290	0.14
	Organization	2.80	1.14	2.98	0.87	−0.99	0.328	0.16
	Grammar	**2.30**	**1.02**	**2.94**	**0.87**	**−4.18**	**0.000**	**0.63**
	Vocabulary	2.62	1.01	2.66	0.85	−0.22	0.826	0.04
	Mechanics	**2.18**	**1.14**	**2.70**	**0.86**	**−2.45**	**0.018**	**0.46**
PEER	**Content**	**2.74**	**1.03**	**3.04**	**0.93**	**−1.85**	**0.071**	**0.29**
	Organization	2.60	1.18	2.78	0.82	−0.98	0.334	0.15
	Grammar	**2.30**	**1.06**	**2.60**	**0.81**	**−1.79**	**0.079**	**0.28**
	Vocabulary	2.24	1.06	2.78	0.76	−3.33	0.002	0.51
	Mechanics	2.34	1.14	2.66	0.80	−1.61	0.114	0.28

Improvements in writing criteria across feedback types

Individual writing criteria analysis revealed a different pattern of improvement for the three feedback types considered. Grammar showed the greatest improvement for students who had the opportunity to receive GPT feedback. A significant increase was seen in the mean scores for grammar which may indicate that students benefited from AI suggestions and explanations to resolve syntactic errors. Content scores were also increasingly significant and GPT feedback may thus have encouraged students to develop their ideas further and work on argument development. Organization and mechanics showed minor improvements that were not statistically significant, while vocabulary was the least improved criterion hinting that GPT feedback might not have been very effective in encouraging lexical diversification.

In contrast, students who received teacher feedback demonstrated the most significant overall improvements across all writing criteria. Content scores increased indicating that expert feedback was particularly effective in enhancing idea development and coherence. Grammar also exhibited a significant increase reinforcing the idea that direct, personalized feedback from an instructor can significantly improve linguistic accuracy. Additionally, mechanics showed a meaningful increase, highlighting the effectiveness of teacher guidance in refining students' understanding of structural and mechanical accuracy. While organization demonstrated a moderate increase, the effect size was smaller, implying that structural improvements may require more extensive iterative revisions. Vocabulary, similar to the other feedback groups, showed minimal progress.

Students who received peer feedback showed noticeable improvements in both content and grammar. Content scores went up suggesting that engaging with peers helped students develop their ideas more fully and think critically about their writing. Grammar also saw significant improvement which may be because giving feedback made students more aware of grammatical errors and how to correct them. On the other hand, improvements in organization and mechanics were minimal and not statistically significant. Vocabulary was the area that showed the least growth, following a similar trend to what was seen in the other feedback categories.

Comparative analysis of total and final score improvements

Table 2b. Results of the students' pre- and post- writing tests across feedback types

Feedback Type	Criterion	Pre M	Pre SD	Post M	Post SD	*t*	*p*	Effect Size
GPT	Total	**25.43**	5.70	**28.19**	4.26	−12.57	.00	.49
	Final Score	**63.58**	14.24	**70.48**	10.65	−12.57	.00	.49
TEACHER	Total	**24.90**	5.23	**27.80**	4.01	−15.66	.00	.56
	Final Score	**62.25**	13.07	**69.50**	10.03	−15.66	.00	.56
PEER	Total	**24.89**	5.20	**28.04**	3.89	−15.51	.00	.61
	Final Score	**62.23**	12.99	**70.10**	9.72	-15.51	.00	.61

Overall scores give a clearer picture of how each type of feedback influenced students' writing progress, as shown in Table 2b. For those who received GPT-generated feedback, total scores and final scores improved significantly. These findings indicate that AI-based feedback had a positive effect, especially when it came to improving structure and grammar.

However, teacher feedback had an even stronger impact. In this group, total scores rose from 24.90 to 27.80 and final scores climbed from 62.25 to 69.50 -- both with $p < 0.001$. The larger effect sizes suggest that working with a teacher helped students not only polish individual elements of their writing but also improve as writers overall.

Peer feedback also made a difference, with both total scores and final scores rising significantly. These results point to the value of collaboration—students seemed to learn a lot through discussion and shared reflection. Still, the slightly smaller improvements compared to teacher feedback suggest that expert input may be more effective for mastering some of the more complex elements of writing.

Generalized Linear Mixed Model (GLMM) Finding

Table 3. Results of Generalized Linear Mixed Model (GLMM)

Source	F	df1	df2	p
Corrected Model	1.675,15	6	143	.000
Pre-Test	3.064,93	1	143	.000
Digital Proficiency	4.671	1	143	.032
Feedback*Digital Proficiency	3.410	2	143	.036

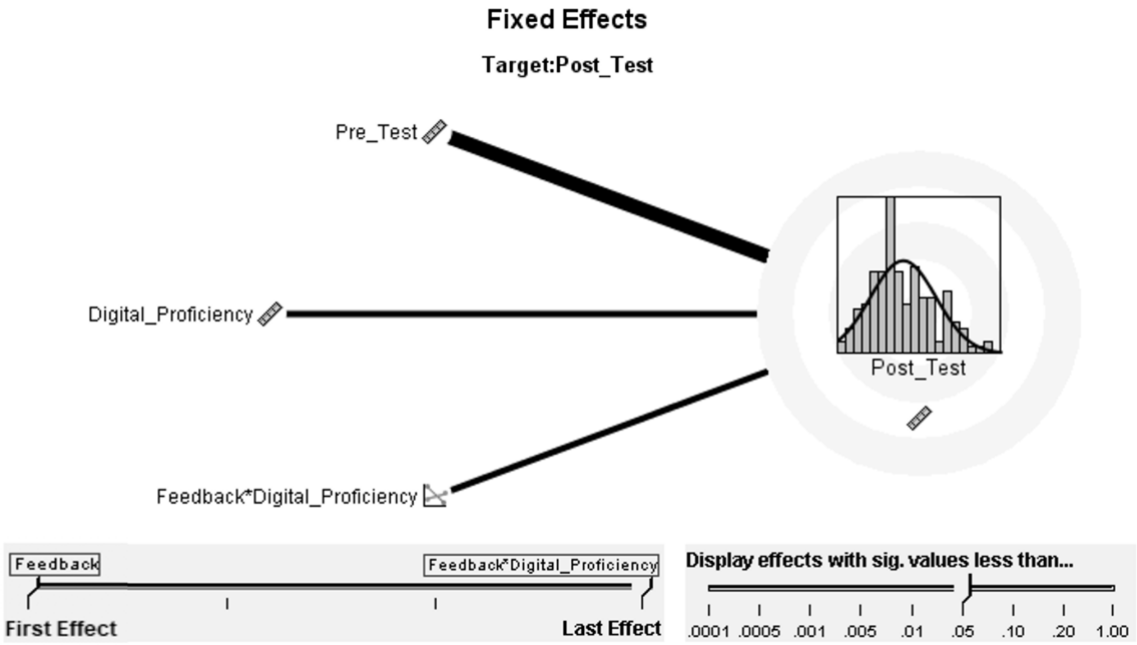

Figure 1 Generalized Linear Mixed Model (GLMM); Fixed Effects

As shown in Table 3 and Figure 1, a Generalized Linear Mixed Model (GLMM) was conducted to examine the interaction between feedback type and digital proficiency on post-test writing scores, while explicitly controlling for students' initial writing performance. This approach incorporates pre-test scores to predict final outcomes. The model was statistically significant ($F(6, 143) = 1{,}675.147$, $p < .001$), demonstrating that the combination of factors effectively accounted for the variance in writing achievement.

The analysis identified the pre-test score as the most powerful predictor of post-test outcomes ($F(1, 143) = 3{,}064.931$, $p < .000$), confirming that baseline ability significantly influences final results. Additionally, students' digital proficiency—treated here as a continuous scale rather than arbitrary categories—exerted a significant main effect on performance ($F(1, 143) = 4.671$, $p = .032$).

Crucially, the interaction between feedback type and digital proficiency was statistically significant ($F(2, 143) = 3.410$, $p = .036$). This finding indicates that the utility of specific feedback modalities, such as AI-generated suggestions, is contingent upon the learner's technological competence. Consequently, these results support the need for differentiated feedback strategies that align with a student's digital readiness to optimize EFL writing development.

Notably, the main effect of feedback type was not statistically significant ($p > .05$), indicating that no single modality—whether AI-generated, teacher-led, or peer-mediated—was universally superior for all students when analyzed in isolation. This lack of a standalone effect suggests that the impact of feedback is not a uniform "treatment" but is instead a contingent process. The non-significance here emphasizes that the primary value of the feedback was not in the modality itself, but in the interaction between the modality and the student's digital readiness.

Calculating Partial Eta-Squared (η_p^2)

Because the Generalized Linear Mixed Model (GLMM) procedure in SPSS does not automatically output this specific metric within the Fixed Effects table, the value was derived manually using the F-statistic and the corresponding degrees of freedom (df1 and df2) provided in the model output.

$$\eta p2 = \frac{F\ x\ df1}{(F\ x\ df1)\ + df2}$$

$$\eta p2 = \frac{3.410\ x\ 2}{(3.410\ x\ 2)\ + 143}$$

$$\eta p2 = \frac{6.82}{149.82} = 0.046$$

An effect size of .046 is interpreted as a small-to-medium effect within the context of EFL writing instruction. This value provides a quantitative measure of the practical significance of the interaction, confirming that the synergy between a student's technological confidence and the type of feedback received has a distinct, measurable impact on their ultimate writing performance.

Independent t-tests: Analysis of Feedback Types and Digital proficiency

Independent t-tests were conducted in examining the differences of means in writing outcomes across digital proficiency groups (High, Moderate, Low) for each of the three feedback types (GPT, Teacher, and Peer). These tests consider the unequal variances and differing sample sizes, so, they are more robust measures of mean differences. Effect sizes (Cohen's d) were also computed to assess the importance, from a practical point of view, of the observed differences.

ChatGPT Feedback

Table 4 Correlation Between Digital Proficiency and Writing Improvement (n = 50)

		Writing (Gain Score)	Digital Proficiency
Writing (Gain Score)	***R*** ***p***	1	.463 .001
Digital Proficiency	***R*** ***p***	.463 .001	1

As shown in Table 4, the analysis revealed a significant positive correlation between students' digital proficiency and their writing gain scores when using AI-generated feedback (r = .463, p = .001). This represents a moderate-to-strong relationship, suggesting that as students' digital proficiency increases, their ability to effectively interpret and implement suggestions from ChatGPT also improves. This supports the finding that AI feedback utility is highly contingent upon the learner's digital proficiency

Teacher Feedback

Table 5. Correlation Between Digital Proficiency and Writing Improvement (n = 50)

		Writing (Gain Score)	Digital Proficiency
Writing (Gain Score)	***R*** ***p***	1	.299 .035
Digital Proficiency	***R*** ***p***	.299 .035	1

Based on Table 5, for the teacher feedback group, a statistically significant but weaker positive correlation was observed (r = .299, p = .035). While higher digital proficiency does provide some advantage—likely due to the digital delivery of teacher comments—the lower correlation coefficient compared to the AI group suggests that teacher feedback is more accessible. It remains effective for a broader range of students regardless of their digital proficiency.

Peer Feedback

Table 6. Correlation Between Digital Proficiency and Writing Improvement (n = 50)

		Writing (Gain Score)	Digital Proficiency
Writing (Gain Score)	***R*** ***p***	1	.284 .046
Digital Proficiency	***R*** ***p***	.284 .046	1

As shown in Table 6, the peer feedback group showed the lowest significant correlation among the three modalities (r = .284, p = .046). This small-to-moderate correlation indicates that while digital proficiency plays a minor role in navigating peer-review platforms, the primary benefits of peer feedback are likely derived from social collaboration and critical awareness rather than digital proficiency. This reinforces the inclusive nature of peer-mediated instruction in digitally diverse environments

Qualitative Findings

After the post-test, we conducted semi-structured interviews with 15 participants who represented a range of digital proficiency, including high, moderate, and low. These interviews were designed to explore students' personal preferences for different types of feedback, how effective they felt each type was, and how easy or difficult they found it to understand and use the feedback they received. The findings indicate that digital proficiency influenced not only preferences but also the perceived utility and accessibility of feedback.

Preferred Feedback Type

There seems to be a close association between participants' feedback preferences and their digital proficiency. Participants with an above-average digital proficiency mostly preferred ChatGPT feedback, citing immediate responses and detailed explanations as reasons. Participant 1, for example, responded, ***"I preferred ChatGPT feedback because it was immediate and gave detailed explanations. I felt it was more consistent compared to human feedback.***" Some participants still preferred teacher feedback, although it was less favored when weighed against AI feedback, on account of the personalized nature of the feedback;

Participant 4 elaborated, ***"I preferred teacher feedback because it provided personalized suggestions."***

Conversely, the digitally moderately competent had more varied tastes. While some liked the convenience and speed of ChatGPT, others favored teacher feedback as being more reliable and consistent. This was seen in the complaint of Participant 6, ***"I liked teacher feedback because it seemed more trustworthy,***" and the negative reverse of Participant 7, who stated**,** ***"I liked peer feedback because it seemed more comparable and less scary."*** For the group of individuals with low digital literacy, the most enjoyed choice was teacher feedback, owing mainly to its personalized and motivational aspect. As Participant 11 summarized, ***"I liked teacher feedback because it was personal and supportive."*** Though rarely cited, peer feedback was sometimes valued for its collaborative and relaxed tone, as Participant 14 explained, ***"Peer feedback was nice because it felt collaborative."***

Most Helpful Feedback Type

Reflecting the difference in perception of feedback usefulness among different digital proficiency levels, participants reflected different ways in which they engaged with feedback. Digital-proficient people most frequently valued ChatGPT-type feedback because it was specific and contained explicit suggestions. Participant 1 reported, ***"ChatGPT was the most helpful because it pointed out grammar mistakes and better ways to phrase things*****."**

However, some weight was still given to the teacher's feedback for more general issues in writing, such as argument clarity: Participant 5 stated, ***"Teacher feedback was helpful for big-picture issues like argument clarity."*** In the intermediate range of digital literacy, most people felt teacher feedback to be most useful because of more subtle points contributed to the holistic approach to improving writing. Participant 8 later explained, ***"Teacher feedback helped me improve the flow and coherence of my essays."*** There were those, however, who found the precision of ChatGPT for technical error detection very useful: Participant 9 added, ***"ChatGPT feedback was the most precise, especially for technical corrections."*** Those considered to have low digital proficiency found teacher feedback as the most useful feedback due to its structured nature and its comprehensive approach. Participant 13 said, ***"Teacher feedback was the most helpful for improving overall quality."*** Less often cited, however, was peer feedback, which was considered advantageous in cultivating an awareness of how my writing affects readers: Participant 12 declared, ***"Peer feedback was the most helpful as it made me realize how my writing affects readers."***

Easiest Feedback Type to Understand and Apply

Easy grasping and application of feedback were critically impacted by digital literacy. Highly digitally proficient participants considered ChatGPT feedback the easiest because it was structured and provided examples. Participant 1 shared that **"*ChatGPT feedback was the easiest to understand because of the clear examples it provided,"*** while Participant 3 added, ***"ChatGPT was straightforward and gave clear suggestions."*** The teacher's feedback was beautifully labeled as the easiest to apply appeal for reasoning behind suggestions mostly by those who considered themselves to have moderate digital skills. Participant 6 explained, ***"Teacher feedback was easier to understand because it explained the reasoning behind the suggestions."***

Nevertheless, some still found the step-by-step instructions offered by ChatGPT especially handy, Participant 7 stating, **"*ChatGPT was easier to apply because it gave step-by-step instructions."*** Conversely, among participants with low digital proficiency, it was always teacher feedback that was seen as the easiest to understand and implement. Participant 15 stressed the clarity of teacher feedback: ***"Teacher feedback was the easiest because it was***

well-structured and clear." Peer feedback was also considered straightforward and accessible; Participant 12 remarked, ***"Peer feedback was easiest because it was less technical*."**

Table 7 below summarizes both performance outcomes and learner engagement across feedback types and digital proficiency levels. It illustrates the interplay between technological confidence and feedback utility.

Table 7. Writing Revision Profiles by Feedback Type and Digital Proficiency Level

Feedback Type	Low Digital Proficiency	High Digital Proficiency
AI-Generated Feedback	Limited uptake; surface-level focus only	High engagement; efficient surface revisions
Teacher Feedback	Reliance on direct instruction	Deep engagement with structural revisions
Peer Feedback	Confusion; inconsistent implementation	Collaborative revisions; mutual validation

Discussion and Implications

This study investigated the comparative effectiveness of AI-generated feedback (ChatGPT), teacher feedback, and peer feedback on Indonesian EFL students' writing performance (RQ1), examined the moderating role of students' digital proficiency (RQ2), and explored their perceptions of and engagement with different feedback types (RQ3). Addressing research gaps highlighted by Hyland and Hyland (2006), Zhang and Hyland (2022), and Escalante et al. (2023), the study responds to calls for deeper analysis of feedback modalities and their pedagogical value. While teacher feedback is often seen as authoritative (Lee, 2019; Cheng & Liu, 2022), peer feedback supports autonomy and collaboration (Farrah, 2012; Yu & Lee, 2016), and AI tools such as ChatGPT show growing potential but remain underexplored in digitally unequal contexts such as Indonesia (Arifin et al., 2024; Mustopa et al., 2024). Grounded in sociocultural theory (Poehner & Leontjev, 2020) and Krashen's Input Hypothesis (Nassaji, 2021), this mixed-methods research positions feedback as a cognitively and socially mediated mechanism in second language writing development. By foregrounding digital proficiency as a moderator, our findings reconceptualize feedback not as a uniform "treatment" but as a contingent tool–learner fit that depends on learners' capacity to interpret, triage, and act on different feedback streams. Pedagogically, this reframing requires designing feedback ecosystems that explicitly teach "feedback literacy" (how to read, evaluate, and apply comments) alongside digital literacy (how to operate and critically use AI tools).

Findings for RQ1 indicate that feedback effectiveness differed by type and writing criterion. Teacher feedback yielded the greatest improvements in grammar, mechanics, and content, aligning with evidence that expert input supports deeper cognitive processing (Li & Zhang, 2022; Waluyo & Apridayani, 2024). Peer feedback was moderately effective, particularly for content and grammar, supporting claims that structured peer interaction enhances critical awareness (Alharbi, 2021; Suherman, 2018). GPT feedback excelled in grammar and organization due to its immediacy and precision (Alghannam, 2024; Chan et al., 2024; Teng, 2024b), but showed limited impact on vocabulary, echoing critiques of its lack of lexical adaptability (Evmenova et al., 2024). These results support that feedback types serve distinct cognitive functions and are most effective when used complementarily (Apridayani & Waluyo, 2025). In practical terms, AI can be positioned earlier in the drafting cycle for rapid, surface-level diagnostics, with teacher and peer feedback reserved for higher-order revision and

rhetorical development; the specific sequencing should be adjusted based on students' digital proficiency and task complexity.

Findings for RQ2 revealed that digital proficiency acted as a significant moderator of feedback effectiveness, particularly for AI-mediated input. The Pearson correlation analysis indicated a moderate-to-strong positive relationship ($r = .463$, $p = .001$) between students' digital proficiency and writing gains in the GPT group. This suggests that the utility of AI-generated suggestions is highly contingent upon a learner's existing digital proficiency, supporting Shen et al. (2023) and Getenet et al. (2024) who link digital self-efficacy to effective use of AI tools . In contrast, the correlations for teacher and peer feedback were statistically significant but notably weaker ($r = .299$ and $r = .284$, respectively), demonstrating that these human-mediated modalities are more digitally inclusive and remain accessible to students regardless of their technical skill levels. These results align with Evans et al. (2010), who conceptualize feedback as contextually mediated rather than uniformly effective. From a sociocultural perspective, meaningful feedback uptake requires both cognitive and digital tools (Apridayani & Waluyo, 2025; Poehner & Leontjev, 2020; Villamil & de Guerrero, 2006), accentuating the risk of widening educational disparities if AI feedback is implemented without addressing digital readiness. Technology integration should include baseline digital diagnostics at course entry, micro-skills training (e.g., interpreting AI comments, spotting hallucinations), and embedded scaffolds within the LMS (templates, exemplars, guardrails) to support equitable uptake. Programmatically, institutions should monitor outcomes by proficiency band, provide targeted access supports (devices, connectivity, offline workflows), and build audit trails of AI use to safeguard transparency and academic integrity.

Qualitative findings for RQ3 revealed that digital proficiency shaped students' trust, emotional engagement, and cognitive preferences regarding feedback. Three key themes, encompassing trustworthiness, cognitive processing, and immediacy, emerged as central to students' judgments. High-proficiency students favored GPT feedback for its speed and clarity, consistent with Marmoah et al. (2024) and Rafida et al. (2024), while moderate-proficiency students showed mixed preferences, balancing usability with perceived reliability. Low-proficiency students preferred teacher feedback for its clarity and motivational support, aligning with Mafulah and Cahyono (2023) and Solak (2024). Peer feedback, though less prominent, was valued for enhancing audience awareness and tone, supporting Palincsar (1998), Chandra et.al (2021), and Damanik (2022). These patterns reflect Carless's (2022) argument that effective feedback engagement depends on both digital fluency and emotional-cognitive readiness. Accordingly, teaching plans should embed affective scaffolds, including norms for constructive dialogue, opportunities for clarification, and scheduled teacher "check-ins", so that lower-proficiency learners are not disadvantaged by the impersonality of automated feedback. Teacher professional development should include AI error-spotting, prompt design aligned to assessment rubrics, and strategies for coaching students to evaluate and selectively adopt AI suggestions.

Overall, the study contributes to the feedback literature by offering a comparative and context-sensitive model that acknowledges both technological and human dimensions of writing instruction. The results advocate for a hybridized feedback approach that blends the authority and contextual relevance of teacher feedback, the collaborative strengths of peer evaluation, and the immediacy and structural precision of AI tools, such as ChatGPT. Crucially, the effectiveness of such integration is contingent upon students' digital proficiency and access, suggesting the need for institutional policies that foster digital competence alongside writing pedagogy. This is especially pressing in Indonesia, where disparities in infrastructure and digital literacy remain significant (Saúde et al., 2024; Achruh et al., 2024). We therefore recommend an implementation package that couples curriculum-embedded digital diagnostics,

staged AI integration by proficiency level, and continuous equity audits of feedback outcomes. Future work should trial these tiered models across semesters, evaluate transfer to independent writing, and examine how sustained "feedback literacy" instruction interacts with digital proficiency to reduce achievement gaps in EFL writing.

Conclusion, Limitations, and Recommendations

This study investigated the relative effectiveness of teacher, peer, and AI-generated feedback on EFL students' writing development and examined how digital proficiency moderated learners' engagement with each feedback type. Teacher feedback produced the most consistent gains in grammar, content development, and mechanics, likely because its specificity and dialogic nature support higher-order reasoning; peer feedback promoted engagement and critical awareness but varied in quality; and AI-generated feedback was especially effective for grammatical accuracy and organizational clarity. Crucially, the effectiveness of AI-generated feedback showed a significant positive correlation with students' digital proficiency, suggesting that the utility of automated suggestions increases linearly with a learner's digital proficiency. Taken together, the findings suggest that feedback modalities are not interchangeable but complementary, and that optimal sequencing should be tailored to the gradient of learners' digital readiness to maximize targeted writing outcomes

Several limitations constrain the generalizability of the findings. First, the single-site context and the predominance of moderately proficient participants limit external validity; outcomes may vary across institutions, learner proficiency profiles, or genre-specific writing demands. Second, the use of intact classes within a quasi-experimental design introduces potential confounds; although prompts, rubrics, and time-on-task were standardized, selection bias and teacher/classroom effects cannot be entirely excluded. Third, possible sources of bias include novelty and Hawthorne effects associated with AI integration, disparities in device access and connectivity, and rater variability in analytic scoring despite rubric calibration. Fourth, the 10-week intervention may be insufficient to assess the durability of learning outcomes, transferability to novel tasks, or the longitudinal development of feedback literacy. Fifth, two subscales measuring digital proficiency demonstrated borderline internal consistency ($\alpha < .70$), and qualitative self-reports remain vulnerable to recall inaccuracies and social desirability bias, necessitating cautious interpretation. Finally, the AI condition was based on a specific model and prompt configuration; ongoing model updates, prompt sensitivity, and current limitations in evaluating discourse-level features present tool-related constraints that may affect replicability and generalization.

Future research should adopt longer, multi-wave longitudinal designs and multi-site or cluster-randomized/stepped-wedge trials to strengthen causal inference and external validity. Priority variables include learner attitudes toward feedback, writing anxiety, feedback literacy, trust in AI, and access/equity indicators, modeled with multilevel approaches to separate student- from class-/teacher-level effects. Process data (e.g., revision logs, keystroke dynamics, LMS clickstream) can illuminate feedback uptake mechanisms, while genre- and discipline-specific tasks can test transferability. Comparative studies across AI models and prompt protocols, alongside cost-effectiveness and integrity audits, would clarify when and how AI should be integrated with teacher and peer feedback to yield durable, equitable gains in EFL writing.

About the Authors

Veni Roza is an ELT scholar specializing in TESOL, grammar, and translation. With a Q1 Scopus-indexed article and multiple Sinta 2 publications, her research advances cross-cultural communication. Previously Head of English Language Education and Secretary of LPM, she

effectively balances high-impact research with dedicated institutional leadership and service. ORCID ID: 0009-0006-5913-6975.

Melyann Melani is an Associate Professor of English at UIN Sjech M. Djamil Djambek Bukittinggi. As Managing Editor of the Modality Journal, she focuses her research on cultural nuances in student translation and pedagogical strategies. Her work is dedicated to enhancing English education and learning outcomes through academic leadership and innovative research. ORCID ID: 0009-0002-0996-4095.

Reflinda is a senior specialist in Applied ELT Pedagogy at UIN Sjech M. Djamil Djambek Bukittinggi. Her research targets classroom interventions, including reading strategies and phonetic analysis. As a central figure in her department's research cluster, she drives academic productivity and uses data-driven feedback to optimize pre-service teacher training programs. ORCID ID: 0009-0007-4278-8288.

Ridha Ahida is a scholar in philosophy and former Rector of UIN Sjech M. Djamil Djambek Bukittinggi. A National Instructor of Religious Moderation, she is globally recognized for her research on interfaith harmony. Her ongoing academic work continues to shape religious discourse and educational excellence through platforms like SINTA and Google Scholar. ORCID ID: 0000-0002-0428-0869.

Rianto is an emerging scholar and lecturer assistant at UIN Sjech M. Djamil Djambek Bukittinggi. Specializing in sociolinguistics and cultural studies, he integrates Islamic narratives into foreign language pedagogy. Through collaborative research on intercultural competence, he leverages faith-based resources to modernize teaching strategies, fostering a unique academic niche aligned with the university's mandate. ORCID ID: 0000-0003-0666-6973.

Safnil Arsyad is a professor of English Education at the Postgraduate Program of the Education Faculty at Bengkulu University in Indonesia. He earned his Master's degree from the University of Canberra and his Ph.D. from the Australian National University. His research interests include Academic Discourse Analysis, English Language Teaching Methods, and Materials Design. ORCID ID: 0000-0003-4174-2556.

Genta Sakti, a senior lecturer at UIN Sjech M. Djamil Djambek Bukittinggi, is an expert in English education methodology. An MA graduate from La Trobe University, his research improves language competency through strategies like Storytelling, THIEVES, and the Flipped Classroom. His work focuses on enhancing reading comprehension and vocabulary mastery within teacher training frameworks.

Yuyud Andam Mellian is a graduate student at UIN Sjech M. Djamil Djambek Bukittinggi, specializing in curriculum theory. Her work explores how global standards, like the Cambridge Curriculum, can be integrated with Indonesia's Merdeka curriculum. Using Culturally-Responsive Teaching (CRT), she advocates for inclusive teaching methods and increased student engagement, as demonstrated by her peer-reviewed research on curriculum implementation.

To Cite this Article

Roza, V., Melani, M., Reflinda, Ahida, R., Rianto, Arsyad, S., Sakti, G. & Mellian, Y. A. (2026). AI vs. human feedback in EFL writing: Examining the interplay between feedback modalities and digital proficiency. *Teaching English as a Second Language Electronic Journal (TESL-EJ), 30*(1). https://doi.org/10.55593/ej.30117int2

References

Achruh, A., Rapi, M., Rusdi, M., & Idris, R. (2024). Challenges and Opportunities of Artificial Intelligence Adoption in Islamic Education in Indonesian Higher Education Institutions. *International Journal of Learning, Teaching and Educational Research*, *23*(11), 423-443. https://doi.org/10.26803/ijlter.23.11.22

Alamäki, A., Khan, U. A., Kauttonen, J., & Schlögl, S. (2024). An Experiment of AI-Based Assessment: Perspectives of Learning Preferences, Benefits, Intention, Technology Affinity, and Trust. *Education Sciences*, *14*(12), 1386. https://doi.org/10.3390/educsci14121386

Alghannam, M. S. M. (2024). Artificial Intelligence as a Provider of Feedback on EFL Student Compositions. *World Journal of English Language*, *15*(2), 161-173. https://doi.org/10.5430/wjel.v15n2p161

Alharbi, M. A. (2021). Students' Perspectives as Providers and Receivers of Peer Formative Feedback on Writing. *World Journal of English Language*, *11*(2), 1-12. https://doi.org/10.5430/wjel.v11n2p1

Allen, D., & Mills, A. (2016). The impact of second language proficiency in dyadic peer feedback. *Language Teaching Research*, *20*(4), 498-513. https://doi.org/10.1177/1362168814561902

Alshumaimeri, Y. A., & Alshememry, A. K. (2023). The extent of AI applications in EFL learning and teaching. *IEEE Transactions on Learning Technologies*, *17*, 653-663. https://doi.org/10.1109/TLT.2023.3322128

AlTwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, *10*(10), 1-12. https://doi.org/10.1016/j.heliyon.2024.e31053

Apridayani, A., & Waluyo, B. (2025). Valuing students' mindsets in essay compositions: Active learning and feedback approaches. *Social Sciences & Humanities Open*, *11*, 101273. https://doi.org/10.1016/j.ssaho.2024.101273

Arifin, M. A., Rahman, A. A., Balla, A., Susanto, A. K., & Pratiwi, A. C. (2024). ChatGPT Affordances and Indonesian EFL Students' Perceptions in L2 Writing: A Collaborative Reflexive Thematic Analysis. *Changing English*, 1-17. https://doi.org/10.1080/1358684X.2024.2418132

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in higher education*, *15*(2), 187-198. https://doi.org/10.1080/13562511003620019

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback?. *International Journal of Educational Technology in Higher Education*, *21*(1), 1-15. https://doi.org/10.1186/s41239-024-00455-4

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

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*(1), 1-18. https://doi.org/10.1186/s41239-023-00411-8

Chan, S. T. S., Lo, N. P. K., & Wong, A. M. H. (2024). Enhancing university level English proficiency with generative AI: Empirical insights into automated feedback and learning outcomes. *Contemporary Educational Technology*, *16*(4), ep541. https://doi.org/10.30935/cedtech/15607

Chandra, R., Yulmiati, Y., & Kemal, E. (2021). An Analysis of Initiation Response Feedback (IRF) Pattern in Classroom Discourse at Half Nelson Movie. *Modality Journal: International Journal of Linguistics and Literature*, *1*(2), 132–144. https://doi.org/10.30983/mj.v1i2.4835

Cheng, X., & Liu, Y. (2022). Student engagement with teacher written feedback: Insights from low-proficiency and high-proficiency L2 learners. *System*, *109*, 102880. https://doi.org/10.1016/j.system.2022.102880

Cuocci, S., Marnani, P. F., Khan, I., & Roberts, S. (2023). A Meta-synthesis of technology-supported peer feedback in ESL/EFL writing classes research: A replication of Chen's study. *Languages*, *8*(2), 1-19. https://doi.org/10.3390/languages8020114

Damanik, J. Y. (2022). Peer Feedback to Improve Indonesian Adult Learners' Writing Skills: A Literature Review. *Journal of English Teaching*, *8*(1), 49-58. https://doi.org/10.33541/jet.v8i1.3253

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, *20*(1), 1-20. https://doi.org/10.1186/s41239-023-00425-2

Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research*, *14*(4), 445-463. https://doi.org/10.1177/1362168810375367

Evenddy, S. S. (2024). Investigating AI's Automated Feedback in English Language Learning. *Foreign Language Instruction Probe*, *3*(1), 76-87. https://doi.org/10.54213/flip.v3i1.401

Evmenova, A. S., Regan, K., Mergen, R., & Hrisseh, R. (2024). Improving writing feedback for struggling writers: Generative AI to the rescue?. *TechTrends*, *68*(4), 790-802. https://doi.org/10.1007/s11528-024-00965-y

Farrah, M. (2012). The Impact of Peer Feedback on Improving the Writing Skills Among Hebron University Students. *An-Najah University Journal for Research-B (Humanities)*, *26*(1), 179-210. https://doi.org/10.35552/0247-026-001-008

Getenet, S., Cantle, R., Redmond, P., & Albion, P. (2024). Students' digital technology attitude, literacy and self-efficacy and their effect on online learning engagement. *International Journal of Educational Technology in Higher Education*, *21*(1), 1-20. https://doi.org/10.1186/s41239-023-00437-y

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218-243. https://doi.org/10.3102/0091732X20903302

Gozali, I., Wijaya, A. R. T., Lie, A., Cahyono, B. Y., & Suryati, N. (2024). Leveraging the potential of ChatGPT as an automated writing evaluation (AWE) tool: Students' feedback literacy development and AWE tools integration framework. *The JALT CALL Journal*, *20*(1), 1-22. https://doi.org/10.29140/jaltcall.v20n1.1200

Hojeij, Z., & Ayber, P. O. (2022). Effectiveness of using digital feedback on EFL student writing skills. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *12*(1), 1-18. https://doi.org/10.4018/IJCALLT.291111

Huang, J., & Teng, M. F. (2025). Peer feedback and ChatGPT-generated feedback on Japanese EFL students' engagement in a foreign language writing context. *Digital Applied Linguistics*, *2*, 102469-102469. https://doi.org/10.29140/dal.v2.102469

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83-101. https://doi.org/10.1017/S0261444806003399

Jansen, T., Höft, L., Bahr, L., Fleckenstein, J., Möller, J., Köller, O., & Meyer, J. (2024). Empirische arbeit: comparing generative AI and expert feedback to students' writing: insights from student teachers. *Psychologie in Erziehung und Unterricht*, *71*(2), 80-92. https://doi.org/10.2378/peu2024.art08d

Jesudass, M., & Mirunalini, M. (2023). Construction and standardisation of digital proficiency scale. *Journal for Re Attach Therapy and Developmental Diversities, 6*(10), 1105-1110. https://jrtdd.com/index.php/journal/article/view/1516

Kartchava, E. (2013). The place of corrective feedback within the major SLA theories. *US-China Foreign Language*, *11*(2), 136-151. https://doi.org/10.17265/1539-8080/2013.02.007

Kusumaningrum, S. R., Cahyono, B. Y., & Prayogo, J. A. (2019). The effect of different types of peer feedback provision on EFL students' writing performance. *International Journal of Instruction*, *12*(1), 213-224. https://doi.org/10.29333/iji.2019.12114a

Lee, I. (2014). Revisiting teacher feedback in EFL writing from sociocultural perspectives. *Tesol Quarterly*, *48*(1), 201-213. https://doi.org/10.1002/tesq.153

Lee, I. (2019). Teacher written corrective feedback: Less is more. *Language Teaching*, *52*(4), 524-536. https://doi.org/10.1017/S0261444819000247

Leow, R. P., & Driver, M. (2021). Cognitive theoretical perspectives of corrective feedback. *The Cambridge handbook of corrective feedback in language learning and teaching*, 65-84. https://doi.org/10.1017/9781108589789.004

Li, D., & Zhang, L. (2022). Contextualizing feedback in L2 writing: the role of teacher scaffolding. *Language awareness*, *31*(3), 328-350. https://doi.org/10.1080/09658416.2021.1931261

Linuwih, E. R., Setiawan, S., & Munir, A. (2024). Student Engagement With Teacher Written Feedback in Online EFL Writing Context. *Theory and Practice in Language Studies*, *14*(7), 2186-2192. https://doi.org/10.17507/tpls.1407.26

Liu, C., & Yu, S. (2022). Reconceptualizing the impact of feedback in second language writing: A multidimensional perspective. *Assessing Writing*, *53*, 100630. https://doi.org/10.1016/j.asw.2022.100630

Lyster, R. (2012). Roles for corrective feedback in second language instruction. *The encyclopedia of applied linguistics*, 1-6. https://doi.org/10.1002/9781405198431.wbeal1028.pub2

Mafulah, S., & Cahyono, B. Y. (2023). Indonesian students' engagement in online EFL writing class and their perceptions on teacher feedback. *Indonesian Journal of Applied Linguistics*, *13*(1), 152-164. https://doi.org/10.17509/ijal.v13i1.58279

Mamad, A., & Vígh, T. (2024). EFL/ESL Teachers' and Students' Perceptions and Practices of Written Feedback in a Higher Education Context. *Profile: Issues in Teachers' Professional Development*, *26*(2), 181-198. https://doi.org/10.15446/profile.v26n2.109141

Marmoah, S., Adika, D., Haryati, S., & Yurni, Y. (2024). Leveraging AI to Optimize English Academic Writing (EAW) in Intelligent Decision Support Systems (IDSS). *Jurnal Ilmiah Universitas Batanghari Jambi*, *24*(2), 1265-1270. https://doi.org/10.33087/jiubj.v24i2.5483

Masrul, M., Erliana, S., Rasyidah, U., & Wicaksono, B. H. (2024). The effect of direct and indirect written corrective feedback on accuracy and fluency of university students' English writing. *World Journal of English Language*, *14*(2), 388-399. https://doi.org/10.5430/wjel.v14n2p388

Mustopa, M., Nasikhin, N., Chamami, R., Nihayah, H., Habibullah, M. R., & Manshur, A. (2024). Challenges in Artificial Intelligence Development in Higher Education in China, India, and Indonesia: International Students' Perspectives. *International Journal of Learning, Teaching and Educational Research*, *23*(2), 354-373. https://doi.org/10.26803/ijlter.23.2.17

Nassaji, H. (2021). Corrective feedback from a sociocultural perspective. *The Cambridge handbook of corrective feedback in language learning and teaching*, 85-107. https://doi.org/10.1017/9781108589789.005

Nodoushan, M. A. S. (2014). Assessing writing: A review of the main trends. *Studies in English Language and Education*, *1*(2), 116–125. https://doi.org/10.24815/siele.v1i2.1831

Oktarina, I. B., Saputri, M. E. E., Magdalena, B., Hastomo, T., & Maximilian, A. (2024). Leveraging ChatGPT to enhance students' writing skills, engagement, and feedback literacy. *Edelweiss Applied Science and Technology*, *8*(4), 2306-2319. https://ideas.repec.org/a/ajp/edwast/v8y2024i4p2306-2319id1600.html

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual review of psychology*, *49*(1), 345-375. https://doi.org/10.1146/annurev.psych.49.1.345

Pang, T. Y., Kootsookos, A., & Cheng, C. T. (2024). Artificial intelligence use in feedback: a qualitative analysis. *Journal of University Teaching and Learning Practice*, *21*(6), 108-125. https://doi.org/10.53761/40wmcj98

Park, S. (2025). Effectiveness and moderating factors of computer-mediated feedback in L2 writing: A meta-analysis. *International Journal of English Language and Literature Studies*, *14*(1), 17-30. https://ideas.repec.org/a/asi/ijells/v14y2025i1p17-30id5328.html

Peungcharoenkun, T., & Waluyo, B. (2024). Students' affective engagements in peer feedback across offline and online English learning environments in Thai higher education. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), 1-22. https://doi.org/10.1186/s40862-024-00286-w

Poehner, M. E., & Leontjev, D. (2020). To correct or to cooperate: Mediational processes and L2 development. *Language Teaching Research*, *24*(3), 295-316. https://doi.org/10.1177/1362168818783212

Rad, H. S., Alipour, R., & Jafarpour, A. (2024). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, *32*(9), 5020-5040. https://doi.org/10.1080/10494820.2023.2208170

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), 987-1016. https://doi.org/10.26811/peuradeun.v12i3.1520

Rahmi, R., Amalina, Z., Andriansyah, A., & Rodgers, A. (2024). Does it really help? Exploring the impact of Al-Generated writing assistant on the students' English writing. *Studies in English Language and Education*, *11*(2), 998-1012. https://doi.org/10.24815/siele.v11i2.35875

Ruwe, T., & Mayweg-Paus, E. (2023). "Your argumentation is good", says the AI vs humans–The role of feedback providers and personalised language for feedback effectiveness. *Computers and Education: Artificial Intelligence*, *5*, 100189. https://doi.org/10.1016/j.caeai.2023.100189

Saúde, S., Barros, J. P., & Almeida, I. (2024). Impacts of Generative Artificial Intelligence in Higher Education: Research Trends and Students' Perceptions. *Social Sciences*, *13*(8), 410. https://doi.org/10.3390/socsci13080410

Shang, H. F. (2022). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*, *30*(1), 4-16. https://doi.org/10.1080/10494820.2019.1629601

Shen, C., Shi, P., Guo, J., Xu, S., & Tian, J. (2023). From process to product: writing engagement and performance of EFL learners under computer-generated feedback instruction. *Frontiers in Psychology*, *14*, 1258286. https://doi.org/10.3389/fpsyg.2023.1258286

Solak, E. (2024). Examining Writing Feedback Dynamics from CHATGPT AI and Human Educators: a Comparative Study. *Педагогика*, *96*(7), 955-970. https://www.ceeol.com/search/article-detail?id=1269098

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, S., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, *91*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Suherman, A. (2018). Exploring EFL Teacher's Interventions on Peer Feedback: a Case Study in a Writing Classroom. *Metathesis: Journal of English Language, Literature, and Teaching*, *2*(2), 194-202. https://doi.org/10.31002/metathesis.v2i2.890

Tan, S., Cho, Y. W., & Xu, W. (2023). Exploring the effects of automated written corrective feedback, computer-mediated peer feedback and their combination mode on EFL learner's writing performance. *Interactive Learning Environments*, *31*(10), 7276-7286. https://doi.org/10.1080/10494820.2022.2066137

Teng, M. F. (2024a). A Systematic Review of ChatGPT for English as a Foreign Language Writing: Opportunities, Challenges, and Recommendations. *International Journal of TESOL Studies, 6*(3), 36-57. https://doi.org/10.58304/ijts.20240304

Teng, M. F. (2024b). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, *7*, 100270. https://doi.org/10.1016/j.caeai.2024.100270

Teng, M. F., & Huang, J. (2025). Incorporating ChatGPT for EFL writing and its effects on writing engagement. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *15*(1), 1-21. https://doi.org/10.4018/IJCALLT.367874

Teng, M. F., & Ma, M. (2024). Assessing metacognition-based student feedback literacy for academic writing. *Assessing Writing*, *59*, 100811. https://doi.org/10.1016/j.asw.2024.100811

Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, *91*, 102247. https://doi.org/10.1016/j.system.2020.102247

Villamil, O. S., & de Guerrero, M. C. (2006). Sociocultural theory: A framework for understanding the social-cognitive dimensions of peer feedback. *Feedback in second language writing: Contexts and issues*, 23-41. https://doi.org/10.1017/CBO9781139524742.004

Wahyudi, D., & Sari, A. (2016). Penggunaan media, variasi, dan umpan balik dalam proses pembelajaran untuk mengoptimalkan hasil belajar siswa. *Jurnal Educative: Journal of Educational Studies*, 1(2), 86-95. https://ejournal.uinbukittinggi.ac.id/index.php/educative/article/view/156/0

Waluyo, B., & Apridayani, A. (2024). Better learning and practice with teacher corrective feedback in higher education: A lesson from Thailand. *Learning: Research and Practice*, *10*(2), 258-275. https://doi.org/10.1080/23735082.2024.2357552

Waluyo, B., & Kusumastuti, S. (2024). Generative AI in student English learning in Thai higher education: more engagement, better outcomes?. *Social Sciences & Humanities Open*, *10*, 101146. https://doi.org/10.1016/j.ssaho.2024.101146

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

Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, *38*(5), 180-210. https://doi.org/10.14742/ajet.7492

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, *49*(4), 461-493. https://doi.org/10.1017/S0261444816000161

Zhang, Z., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, *51*, 100586. https://doi.org/10.1016/j.asw.2021.100586

Zhang, Z., & Hyland, K. (2023). The role of digital literacy in student engagement with automated writing evaluation (AWE) feedback on second language writing. *Computer Assisted Language Learning*, 1-26. https://doi.org/10.1080/09588221.2023.2256815

Zou, D., Xie, H., & Wang, F. L. (2023). Effects of technology enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. *Journal of Computing in Higher Education*, *35*(1), 166-185. https://doi.org/10.1007/s12528-022-09337-y