BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Revising with Intelligence: ChatGPT Feedback and Its Impact on EFL Students' Revision and Self-Efficacy

Yang Yang , ***Supyan Hussin**** , ***Harwati Hashim***

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

ABSTRACT

While prior research on AI-assisted writing has primarily focused on surface-level gains in accuracy or anxiety reduction, this study investigates how ChatGPT-supported feedback affects EFL learners' writing development across behavioral, cognitive, and affective dimensions. Employing a quasi-experimental mixed-methods design, 82 university-level EFL students were assigned to either a ChatGPT-supported or teacher-supported writing group. Both groups completed three writing tasks over ten weeks. The ChatGPT group received only ChatGPT feedback, while the teacher group received conventional teacher comments. Pre- and post-intervention measures included writing self-efficacy questionnaires, writing samples, and semi-structured interviews. Quantitative results showed that ChatGPT-supported feedback significantly increased revision productivity and led to more macro-level and content-based revisions. Self-efficacy gains in the ChatGPT group were also significantly higher across all three measured dimensions: substantive revision, discourse organization, and writing conventions. Qualitative findings revealed that students in the ChatGPT group perceived the feedback as clearer, more actionable, and more dialogic, which promoted revision ownership, strategic engagement, and confidence. These findings suggest that ChatGPT can function as a cognitive and emotional scaffold, enabling deeper learner engagement with feedback and supporting self-regulated revision. In contrast, teacher feedback, while valued for its credibility, was perceived as less interactive and less helpful for content-level improvement. This study highlights the pedagogical potential of integrating generative AI into process-based writing instruction. By fostering greater revision productivity, deeper feedback engagement, and enhanced writing self-efficacy, ChatGPT can serve as a catalyst for developing more autonomous and reflective EFL writers.

Keywords: ChatGPT; EFL Writing; Feedback Engagement; Revision Practices; Writing Self-Efficacy; Generative AI; Mixed-Methods Research

*CORRESPONDING AUTHOR:

Supayan Hissin, Faculty of Education, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia; Email: supyan@ukm.edu.my

ARTICLE INFO

Received:3 May 2025 | Revised: 3 June 2025 | Accepted: 10 June 2025 | Published Online: 15 July 2025

DOI: https://doi.org/10.30564/fls.v7i7.9845

CITATION

Yang, Y., Hussin, S., Hashim, H., 2025. Revising with Intelligence: ChatGPT Feedback and Its Impact on EFL Students' Revision and Self-Efficacy. Forum for Linguistic Studies. 7(7): 352–367. DOI: https://doi.org/10.30564/fls.v7i7.9845

1. Introduction

Revision is a defining component of EFL writing development, functioning as a critical indicator of how learners process and act upon feedback [1]. It captures not only the frequency of textual changes but also the depth of those changes—from surface-level edits (e.g., grammar, mechanics) to meaning-level transformations such as rhetorical restructuring, discourse reorganization, or ideational elaboration [2]. While revision has long been studied as a cognitive and rhetorical activity [3], more recent research also frames it as an observable marker of engagement with feedback, indicating both behavioral initiative and cognitive investment [4]. Understanding the extent, manner, and stimulus of learners' revision provides valuable insight into the effectiveness of different feedback approaches.

Revision is fundamentally shaped by the nature, clarity, and responsiveness of the feedback students receive. Effective feedback guides learners toward improved clarity, coherence, and accuracy in writing, while also fostering deeper cognitive and motivational engagement [5]. In many L2 classrooms, however, feedback remains narrowly focused on linguistic accuracy and is frequently non-interactive, lacking the responsiveness and dialogic scaffolding now shown to support learner autonomy and self-regulated revision strategies [6].

These limitations have paved the way for alternative forms of feedback, most notably through AI-powered tools like ChatGPT. ChatGPT provides immediate, elaborated, and interactive responses, simulating a conversational partner rather than a static evaluator. This dialogic affordance enables learners to engage in revision through iterative questioning, clarification, and testing, all of which align well with process-oriented writing pedagogy. Nevertheless, much of the current research has concentrated on surface-level linguistic outcomes, overlooking how AI-generated feedback influences learners' broader revision behaviors, such as structural changes and ideational development [7].

Recent scholarship proposes that feedback engagement should be understood as a multidimensional process involving behavioral, cognitive, and affective components [8]. These dimensions have become especially salient in AI-mediated writing contexts, where learners' uptake and emotional regulation are closely linked to revision strategies and confidence [9]. This conceptualization is grounded in established learning theories: sociocognitive theory highlights the role of feedback in building self-efficacy and promoting self-regulated revision strategies while sociocultural theory emphasizes the importance of interactive support within a learner's zone of proximal development [10,11]. Within these theoretical perspectives, ChatGPT serves as a dialogic scaffold that fosters ideational experimentation and deeper textual control [12].

Despite these theoretical foundations, empirical research has yet to systematically investigate how AI-generated feedback influences learners across the behavioral, cognitive, and affective domains. To address this gap, the present study investigates how ChatGPT-supported feedback influences EFL learners' writing development across three interrelated domains of engagement. Specifically, it explores the following research questions:

1. How does revision productivity differ between ChatGPT-supported and teacher-supported EFL writing groups across tasks?
2. What are the differences in the distribution of formal, meaning-preserving, microstructural, and macrostructural revisions between ChatGPT-supported and teacher-supported writing groups?
3. How does feedback engagement through ChatGPT affect students' writing self-efficacy?
4. How do EFL students perceive and cognitively engage with feedback in ChatGPT-supported versus teacher-supported environments?

2. Literature Review

2.1. Revision as a Window into Feedback Engagement

In EFL writing research, revision is widely recognized as a cognitively demanding and developmentally significant process that reflects learners' engagement with feedback [13]. Rather than being limited to surface-level correction, revision often reveals learners' interpretation, prioritization, and internalization of feedback. The distinction between surface (e.g., grammar, word choice) and meaning-level revisions (e.g., organization, argumentation) is foundational in feedback studies and informs recent research on feedback literacy and writing engagement [14].

More recent research highlights revision as an observable indicator of behavioral engagement with feedback. Meaningful revisions, especially those that reflect structural or rhetorical reworking, often indicate higher levels of cognitive involvement and learner agency. The clarity and specificity of feedback critically shape students' revision behaviors [15]. Vague or evaluative comments may lead to superficial edits, while directive and elaborated feedback can foster deeper revisions. The distinction between surface- and meaning-level revisions also serves as a reliable behavioral proxy for the degree to which learners cognitively process and internalize feedback [6].

2.2. Feedback Modalities and the Emergence of AI Support

The nature of feedback delivery plays a central role in shaping learners' revision behavior and engagement. Traditional teacher feedback, while pedagogically informed and context-sensitive, has been critiqued for its limited immediacy, minimal elaboration, and one-way communication style [7]. These limitations may hinder learners' ability to act on feedback or develop independent revision strategies, particularly when feedback is delayed or lacks elaboration, thereby reducing opportunities for reflection and iterative revision.

By contrast, large language models such as ChatGPT have introduced a new paradigm for writing feedback, offering instant, elaborated, and often dialogic responses [16]. These affordances may increase learner autonomy and metacognitive engagement by allowing students to test, revise, and iterate their writing in real time. Some studies suggest that the immediacy and personalization of ChatGPT-generated feedback contribute to reduced anxiety, enhanced motivation, and deeper revisions. However, few studies have directly compared ChatGPT feedback with teacher-generated feedback in terms of revision productivity, revision type, and learner perception [7].

2.3. Theoretical Framework: Feedback Engagement and Writing Self-Efficacy

Feedback engagement has been proposed to encompass behavioral, cognitive, and affective components. Behavioral engagement refers to how students act on feedback—through revision frequency, range, and depth; cognitive engagement involves learners' interpretation, decision-making, and strategy use; and affective engagement includes emotional responses, confidence, and writing beliefs, all of which interact dynamically during revision processes [17]. These dimensions are theoretically grounded in both sociocognitive and sociocultural perspectives [18]. Sociocognitive theory emphasizes the centrality of self-efficacy in regulating behavior and sustaining persistence in challenging tasks, a principle that remains foundational in contemporary L2 writing studies. Recent findings suggest that learners with stronger writing self-efficacy are more likely to pursue meaning-level revisions, persist through complex feedback and stimulate sustained revision behaviors, though the underlying affective pathways remain a promising area for further research [19,20]. Sociocultural theory complements this view by foregrounding the role of mediated scaffolding. Empirical work illustrates that AI-generated feedback tools like ChatGPT can activate ZPD processes by prompting iterative exploration and offering adaptive linguistic support [21].

3. Research Methodology

3.1. Research Design

This study adopted a quasi-experimental mixed-methods design to examine how different feedback sources influence EFL writing development [22]. Two intact classes from a Chinese public university were assigned to the experimental group (ChatGPT-supported feedback) and the control group (teacher feedback). Both groups followed the same syllabus, instructional schedule, and teaching materials; feedback source was the only manipulated variable. The 11-week intervention included three writing tasks, each involving drafting, feedback, and revision. Writing self-efficacy was measured through pre- and post-intervention questionnaires. To capture a more holistic picture of student experience, semi-structured interviews were conducted to explore perceptions of feedback clarity, usefulness, and emotional response [23].

Although random assignment was not feasible, baseline equivalence was ensured through pretest analysis of language errors and feedback coverage by expert raters. This quasi-experimental design preserved ecological validity by maintaining authentic instructional settings while still allowing for controlled comparison [24]. Integrating quantitative indicators with learner perceptions enabled the study to probe not only what changed in students' writing but also how they processed and acted upon feedback cognitively and affectively [22].

3.2. Participants and Setting

This study was conducted at an undergraduate university in China, involving 82 first-year students enrolled in a required course titled *College English Writing I*. All were non-English majors with intermediate proficiency, approximately aligned with CEFR B1 and approaching the CET-4 benchmark. A convenience sampling strategy was adopted, with two intact classes selected for accessibility and instructional feasibility. These classes were assigned to the experimental and control groups, respectively. First-year students were chosen as their extrinsically motivated learning context provided a relevant setting to evaluate AI-mediated feedback. One student in the experimental group failed to complete the pretest, and one in the control group was excluded due to prolonged absence, resulting in 40 students in the experimental group and 42 in the control group.

Instructional content, pacing, and materials were held constant across both groups. All students followed the same syllabus and textbook, and the course was delivered by the same instructor. A university-supported online learning platform (Chaoxing) was used to distribute materials and collect assignments. Ethical approval was obtained, and all participants signed informed consent forms. Anonymity and confidentiality were maintained throughout the study.

3.3. Instruments

This study employed three primary instruments: a writing self-efficacy questionnaire, a set of writing tasks, and a semi-structured interview protocol.

3.3.1. Writing Self-Efficacy Questionnaire

The questionnaire was adapted from a validated instrument developed by Zhan and Yan [25], which demonstrated excellent internal consistency (Cronbach's α = 0.946–0.963 across subscales). Its construct validity was confirmed via exploratory and confirmatory factor analyses, supporting a three-factor structure (SRE, DSS, and WCS) with strong model fit (CFI = 0.987, TLI = 0.985, RMSEA = 0.042). Discriminant and criterion validity were also established through significant correlations with writing motivation and task performance.

3.3.2. Writing Tasks

Prior to the intervention, all participants completed a pre-writing task designed to establish a baseline for later comparisons. In this task, students responded to the prompt, "Should college education be free for all students?" by writing a short argumentative essay of approximately 150–200 words. The writing was completed individually under exam-like conditions, with a 30-minute time limit and no access to AI tools or peer support. These samples served to assess initial writing proficiency and to track subsequent changes in revision behavior.

Three additional writing tasks were administered during the intervention phase, each incorporating either ChatGPT- or teacher-supported feedback and designed to elicit varied rhetorical structures and revision strategies:

Task 1: Should college students take part-time jobs or focus on their studies?

Task 2: Why do young people prefer online shopping, and what are the effects?

Task 3: What are the impacts of excessive social media use, and how can it be addressed?

3.3.3. Semi-Structured Interview Protocol

A semi-structured interview protocol (see **Appendix A**) was used to explore learners' perceptions of feedback. Interview questions addressed clarity, perceived usefulness, emotional responses, and revision decision-making. Interviews were conducted individually in the participant's preferred language and audio-recorded with consent. Pseudonyms were used to protect confidentiality and reduce response bias [12].

3.4. Feedback Delivery Protocol

To ensure comparability between feedback conditions, this study adopted structured procedures covering feedback design, delivery, and monitoring.

In the teacher-supported group, feedback was provided by an instructor holding a Master's degree in Applied Linguistics with over eight years of experience teaching university-level English writing. Feedback was delivered in two stages: (1) written annotations on student drafts focusing on organization, clarity, coherence, and language use; and (2) brief in-class clarification sessions (3–5 minutes per student) conducted in small groups. Written comments were in English, while oral explanations were primarily in Mandarin to ensure accessibility.

In the ChatGPT-supported group, students received brief training on effective prompt usage and were instructed to use the standardized prompt to elicit balanced and pedagogically aligned feedback:

> *"You are an English writing teacher. Please read the student's essay below and provide clear, specific revision suggestions in bullet points. Focus on grammar and word use, clarity of meaning, paragraph structure, and overall coherence. Do not rewrite the essay."*

Although AI-generated feedback was in English, students were encouraged to engage in follow-up queries (e.g., asking for word meanings or clarification in either Chinese or English) within the same interface.

All writing tasks were completed in class under teacher supervision. Students were explicitly instructed that ChatGPT was to be used solely for revision feedback, not for full-text generation which is in alignment with institutional AI-use policies. These rules were reinforced at the start of each session. To ensure compliance, students submitted screenshots of their ChatGPT interactions, which were reviewed prior to data analysis. This combination of real-time supervision and artifact verification effectively minimized the risk of unauthorized AI-generated writing and ensured that all submitted drafts reflected students' genuine revision efforts.

3.5. Data Collection and Analysis

Data were collected over an 11-week period, comprising three phases: a pretest phase, an instructional intervention phase, and a posttest phase. In Week 1, all participants completed a writing self-efficacy questionnaire and a baseline writing task, which was used solely for expert-coded comparisons of initial language error load and feedback coverage. During Weeks 2–10, students completed three writing tasks, each structured around drafting, receiving feedback, and revising. The only procedural difference between groups lay in the source of feedback: the experimental group received AI-generated comments via ChatGPT, while the comparison group received conventional teacher feedback. In Week 11, all participants completed the post-intervention self-efficacy questionnaire. The detailed timeline and group-specific procedures are summarized in **Table 1**.

Table 1. Study Timeline and Group Procedures.

Week(s)	Activity	ChatGPT Group	Teacher Group
Week 1	Pre-test	Self-efficacy questionnaire Writing-pretest	Self-efficacy questionnaire Writing-pretest
Weeks 2–4	Task 1	ChatGPT feedback → Revision	Teacher feedback → Revision
Weeks 5–7	Task 2	ChatGPT feedback → Revision	Teacher feedback → Revision
Weeks 8–10	Task 3	ChatGPT feedback → Revision	Teacher feedback → Revision
Week 11	Post-test	Self-efficacy questionnaire + Interview (selected)	Self-efficacy questionnaire + Interview (selected)

Note. Both groups received identical instructions and writing tasks; the only difference lay in the feedback source and delivery mode.

3.6. Data Collection

Writing samples collected across the study served distinct analytical purposes. To establish baseline comparability before the intervention, four expert raters independently evaluated students' pretest writing samples and the feedback provided by either ChatGPT or teachers. They coded the number of language errors and calculated the proportion of those errors addressed in the feedback (feedback coverage). This preliminary coding ensured that the two groups started from similar levels of writing accuracy and feedback exposure; detailed statistical comparisons are reported in the Results section. Inter-rater reliability for this stage was high (Cohen's κ = 0.87), indicating strong consistency in error and feedback coding.

During the intervention phase, writing samples from the three subsequent tasks were collected to examine revision behaviors. For RQ1, revision productivity was calculated based on the total number of revisions students made after receiving feedback. For RQ2, all revisions were coded into surface-level (e.g., grammar, punctuation) and content-level (e.g., argumentation, structure) categories. This classification enabled cross-group comparison of revision patterns. The same group of expert raters conducted this revision coding (Cohen's κ = 0.81), ensuring the validity of the categorization process.

Self-efficacy questionnaires administered at the beginning and end of the study measured changes in students' perceived writing competence, their confidence in handling writing tasks, and their resilience in facing writing challenges (RQ3). To deepen the understanding of feedback perception and engagement (RQ4), Semi-structured interviews were conducted with 14 participants (7 from each group), selected through purposive sampling to ensure variation in writing performance and gender. This approach aimed to gather diverse perspectives on feedback engagement and provide qualitative depth to complement the quantitative results.

3.7. Data Analysis

To examine the effects of different feedback sources on students' writing development, the analysis combined non-parametric and parametric techniques tailored to the nature of the data. Given the non-normal distribution and ordinal characteristics of revision-related data, Mann-Whitney U tests were applied to compare revision productivity and types across groups, which offers robust insight into how students responded behaviorally to feedback. In contrast, changes in writing self-efficacy were analyzed using repeated-measures ANOVA, which enabled the identification of both within-group progress and between-group differences over time [26].

To enrich the interpretation of these quantitative trends, thematic analysis was conducted on interview data to explore students' perceptions of feedback clarity, responsiveness, and cognitive engagement. Following Braun and Clarke's established framework [27], this method facilitated the identification of patterns in how learners made sense of and acted upon the feedback they received. The combination of statistical analysis and qualitative interpretation reflects a pragmatic mixed-methods approach [22], allowing the study to capture not just what changed, but how and why those changes occurred.

4. Results

Before addressing the research questions, a baseline comparison was conducted to ensure group equivalence in terms of initial writing accuracy and feedback input. Four expert raters independently evaluated the pretest writing samples and the corresponding feedback provided to each group. The results revealed no statistically significant difference in the average number of language-related errors between the ChatGPT-supported group (M = 15, SD = 3) and the teacher-supported group (M = 14, SD = 3). Similarly, the proportion of those errors addressed in the feedback was comparable across groups (ChatGPT: 93.0%; Teacher: 92.1%). These findings suggest that both groups began the intervention phase from a comparable starting point in terms of language accuracy demands and feedback input, providing a valid foundation for subsequent analyses of revision outcomes and learner engagement under differing feedback conditions.

4.1. Comparing Revision Productivity across Feedback Conditions

Following the pretest comparisons, students' revision productivity across the three writing tasks was analyzed to address RQ1. The ChatGPT-supported group consistently outperformed the teacher-supported group in revision productivity across all three writing tasks. As shown in **Table 2**, the average number of revisions in the ChatGPT group increased steadily from Task 1 (M = 11.55, SD = 4.11) to Task 3 (M = 13.55, SD = 4.11), compared to the teacher group's smaller gains (Task 1: M = 7.67, SD = 3.21; Task 3: M = 9.67, SD = 3.21). Mann–Whitney U tests revealed statistically significant differences at each task stage (Task 1: U = 387.5, $p < 0.001$; Task 2: U = 356.0, $p < 0.001$; Task 3: U = 340.0, $p < 0.001$), confirming that ChatGPT feedback led to significantly more frequent revisions. These results showed consistent differences in revision frequency between groups across all tasks.

Figure 1 illustrates the distribution of revision counts across tasks and groups. The ChatGPT-supported group not only had higher median revision counts in each task, but also showed a wider interquartile range and higher upper whiskers—indicating both greater average productivity and more high-performing outliers. In contrast, the teacher-supported group displayed a more compressed distribution, with lower medians and fewer instances of extensive revision. These visual patterns reinforce the quantitative findings and suggest that ChatGPT feedback encouraged wider distribution and greater output of revisions in the ChatGPT group across all tasks.

Table 2. Revision Productivity by Group Across Writing Tasks.

Task	Mean (SD) (Teacher, n = 42)	Mean (SD) (ChatGPT, n = 40)	U	p-value	Significance
Task 1	7.67 (3.21)	11.55 (4.11)	387.5	0.00003	***
Task 2	8.88 (3.20)	13.2 (4.27)	356	0.00001	***
Task 3	10.26 (3.21)	14.82 (4.38)	340	0	***

Note. Mann–Whitney U tests were conducted to compare revision productivity between the teacher-supported and ChatGPT-supported groups across three writing tasks. Asterisks indicate levels of statistical significance ($p < 0.05$, *$p < 0.01$, *$p < 0.001$).

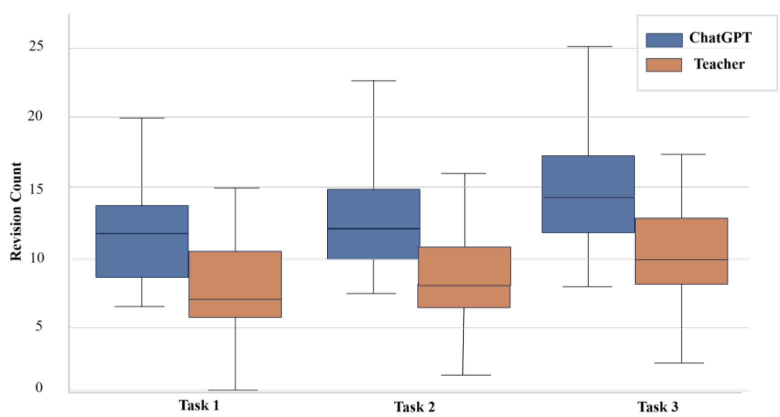


Figure 1. Distribution of Revision Counts by Group Across Writing Tasks.

4.2. Comparing Revision Types across Feedback Conditions

To address RQ2, Mann-Whitney U tests were conducted to compare the frequency of revision subtypes between the ChatGPT-supported and teacher-supported groups, as the data violated the assumption of normality. As shown in **Table 3**, the ChatGPT-supported group consistently produced significantly more text-base changes (namely microstructure and macrostructure revisions) than the teacher-supported group (all $p < 0.001$), with large effect sizes (r = 0.50–0.72). Notably, macrostructure changes showed the largest and most consistent group gap, with effect sizes ranging from $r = 0.41$ to 0.46, indicating a moderate to large effect. In Task 3, for instance, the ChatGPT-supported group averaged 5.7 macrostructure changes, compared to only 1.3 in the teacher-supported group.

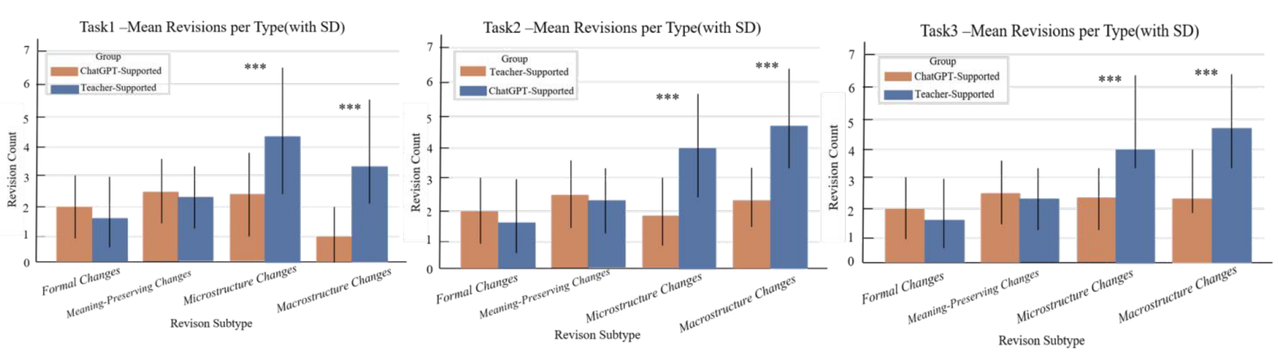
Similarly, microstructure revisions also differed significantly (e.g., Task 3: $U = 424.5$, $p < 0.001$), with the ChatGPT group producing a higher number of local content refinements. These meaning-oriented changes, involving deeper restructuring of sentence-level and discourse-level content, showed a progressive increase in the ChatGPT group across tasks. By contrast, no significant differences were observed in formal or meaning-preserving changes between the two groups (all $p > 0.05$), highlighting that the main divergence lies in how learners revised beyond surface-level concerns.

This quantitative trend is visually reinforced by the bar chart (**Figure 2**), the ChatGPT-supported group demonstrated a consistently higher count of content-oriented revisions—particularly microstructure and macrostructure changes—across all tasks (**Figure 2(A)–(C)**). The differences became more pronounced in Tasks 2 and 3 (**Figure 2(B)** and **(C)**), where macrostructure changes in the ChatGPT group surpassed 5 per task, compared to fewer than 3 in the teacher group. These trends suggest not only greater engagement with feedback, but also a gradual deepening of revision sophistication under ChatGPT-supported conditions.

Table 3. Revision Productivity by Group Across Writing Tasks.

Task	Revision Type	Teacher Group M (SD)	ChatGPT Group M (SD)	U	*p*	Effect Size (r)
1	Formal Changes	2.0 (1.10)	1.7 (1.14)	951	0.287	0.12
	Meaning Preserving Changes	2.4 (1.17)	2.2 (1.11)	921	0.433	0.09
	Microstructure Changes	2.3 (1.56)	4.4 (1.98)	359	***	0.5
	Macrostructure Changes	1.0 (1.00)	3.3 (2.08)	275	***	0.59
2	Formal Changes	2.0 (1.10)	1.7 (1.14)	951	0.287	0.12
	Meaning Preserving Changes	2.4 (1.17)	2.2 (1.11)	921	0.433	0.09
	Microstructure Changes	1.9 (1.09)	4.0 (1.57)	239	***	0.63
	Macrostructure Changes	2.4 (0.99)	4.7 (1.53)	165	***	0.7
3	Formal Changes	2.0 (1.10)	1.7 (1.14)	951	0.287	0.12
	Meaning Preserving Changes	2.4 (1.17)	2.2 (1.11)	921	0.433	0.09
	Microstructure Changes	2.4 (0.99)	4.7 (1.53)	165	***	0.7
	Macrostructure Changes	2.9 (1.09)	5.0 (1.57)	239	***	0.63

Note. Mann–Whitney U tests were conducted to compare revision subtypes between the ChatGPT-supported and teacher-supported groups. Significance levels: $p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***.



(A) **(B)** **(C)**

Figure 2. Mean Revision Counts by Subtype, Task, and Group. **(A)** Task 1. **(B)** Rask 2. **(C)** Task 3.

Note. Error bars = SD. Asterisks denote significant differences (***$p < 0.001$).

4.3. RQ3 Impact of Feedback Conditions on Writing Self-Efficacy

Table 4 presents descriptive statistics of students' writing self-efficacy across three dimensions—substantive revision efficacy (SRE), discourse synthesis efficacy (DSS), and writing conventions self-efficacy (WCS)—for both the teacher-supported and ChatGPT-supported groups. At pretest, the two groups demonstrated comparable self-efficacy levels across all dimensions, with mean scores ranging from 3.08 to 3.36. However, by the posttest, the ChatGPT-supported group exhibited substantially greater gains, particularly in SRE (from 3.34 to 3.97) and WCS (from 3.30 to 4.00), compared to relatively modest improvements in the teacher-supported group. The more pronounced changes in the ChatGPT group suggest a stronger sense of writing competence and confidence, especially in revising content and applying conventions. These descriptive trends provide preliminary support for the hypothesis that AI-mediated feedback fosters enhanced self-perception in writing tasks. To confirm the significance of these changes and group differences, inferential analysis was subsequently conducted, as reported in **Table 4**.

All self-efficacy variables met the assumption of normality across groups and time points. A mixed-design ANOVA revealed significant time × group interaction effects across all three dimensions (**Table 5**): SRE and DSS each showed $F(1, 80) = 10.55$, $p = 0.002$, $\eta^2 = 0.117$; WCS exhibited a stronger interaction, $F(1, 80) = 27.29$, $p < 0.001$, $\eta^2 = 0.254$. Main effects of time and group were also significant across the board (*ps* < 0.001), with large effect sizes (η^2 ranging from 0.447 to 0.668).

Table 4. Descriptive Statistics of Writing Self-Efficacy by Group and Time.

Group		**Pre_Test**			**Post_Test**		
		Pre_SRE	**Pre_DSS**	**Pre_WCS**	**Post_SRE**	**Post_DSS**	**Post_WCS**
Teacher-Supported (n=42)	Mean	3.13	3.083	3.16	3.41	3.41	3.41
	SD	0.3	0.2219	0.315	0.303	0.248	0.264
ChatGPT-Supported (n=40)	Mean	3.34	3.363	3.3	3.97	3.92	4
	SD	0.268	0.218	0.24	0.223	0.197	0.296

Note. SRE = Substantive Revision Efficacy; DSS = Discourse Synthesis Self-Efficacy; WCS = Writing Conventions Self-Efficacy.

Table 5. Results of Two-Way Repeated Measures ANOVA on Writing Self-Efficacy Scores.

Dimension	**Effect**	**F**	***P***	**Partial η²**
SRE	Time	160.66	< 0.001	0.668
	Time × Group	10.55	0.002	0.117
DSS	Time	160.66	< 0.001	0.668
	Time × Group	10.55	0.002	0.117
WCS	Time	124.05	< 0.001	0.608
	Time × Group	27.29	< 0.001	0.254

Note. SRE = Substantive Revision Efficacy; DSS = Discourse Synthesis Self-Efficacy; WCS = Writing Conventions Self-Efficacy. Significant interactions (Time × Group) suggest greater improvement in the ChatGPT-supported group across all three dimensions.

To further assess practical significance, Cohen's *d* was calculated. The ChatGPT group showed very large within-group improvements: SRE ($d = 2.56$), DSS ($d = 2.68$), and WCS ($d = 2.60$), compared to medium-to-large effects in the teacher group: SRE ($d = 0.93$), DSS ($d = 1.39$), and WCS ($d = 0.86$). Between-group post-test comparisons also yielded large effects (SRE $d = 2.10$; DSS $d = 2.27$; WCS $d = 2.11$), confirming the superior impact of ChatGPT-supported feedback on students' self-efficacy. Although the ChatGPT group exhibited slightly higher pre-test means across dimensions, these differences were not statistically significant. The substantially greater post-test gains and effect sizes in this group are therefore attributable to the intervention rather than baseline advantage.

To further interpret the significant Time × Group interaction effects identified in the mixed-design ANOVA, appropriate post-hoc analyses were conducted. Paired-sample t-tests showed significant pre-to-post gains in writing self-efficacy across all three dimensions within both groups (ps < 0.001), with large effect sizes (Cohen's $d = 0.80–2.55$). Independent-sample t-tests comparing post-test scores revealed that the teacher-supported group outperformed the ChatGPT-supported group on all dimensions: Substantive Revision Efficacy ($t(\approx 80) = –9.50$, $p < 0.001$, $d = 2.08$), Discourse Synthesis Self-Efficacy ($t(\approx 80) = –10.24$, $p < 0.001$, $d = 2.25$), and Writing Conventions Self-Efficacy ($t(\approx 80) = –9.43$, $p < 0.001$, $d = 2.09$). As summarized in **Table 6**, these results confirm that while both feedback types enhanced self-efficacy, teacher-supported feedback yielded significantly greater post-intervention gains.

Table 6. Post-Hoc Comparisons of Writing Self-Efficacy Scores between Groups at Post-Test.

Dimension	ChatGPT Mean	Teacher Mean	t	*p*	Cohen's d
SRE	3.41	3.97	–9.5	< 0.001	2.08
DSS	3.41	3.92	–10.24	< 0.001	2.25
WCS	3.41	4	–9.43	< 0.001	2.09

Note. Post-test scores were compared between the ChatGPT- and teacher-supported groups using independent-sample t-tests. Cohen's *d* represents effect size for group differences.

Figure 3 visualizes the estimated marginal means of writing self-efficacy across time and groups for the three subdimensions: self-regulation efficacy (SRE), declarative strategy self-efficacy (DSS), and writing confidence/self-expression (WCS). In all three graphs, a visible interaction pattern is evident: both groups showed improvement from pre-test (Time 1) to post-test (Time 2), but the increase was consistently more substantial in the ChatGPT-supported group. The gap between the two groups widened over time, particularly in the WCS dimension, where the post-test mean for the ChatGPT group approached 4.0, compared to just above 3.4 in the teacher-supported group. The error bars (95% CI) do not overlap in post-test estimates for WCS, indicating a robust group difference. The upward trajectory in the ChatGPT group across all three dimensions supports the statistical findings of significant time × group interaction effects.

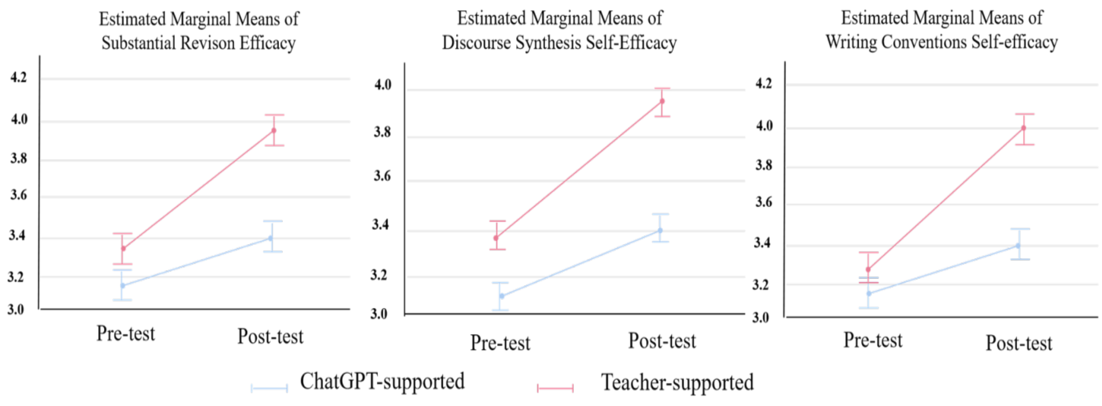


Figure 3. Changes in Writing Self-Efficacy Across Time by Group, with 95% Confidence Intervals.

4.4. Cognitive Engagement with Feedback: Insights from Student Interviews

Drawing on thematic analysis of semi-structured interviews, this section explores students' cognitive engagement with feedback under two distinct conditions: ChatGPT-supported and teacher-supported writing. While participants in both groups valued feedback as integral to their writing development, ChatGPT-supported students exhibited deeper cognitive engagement, characterized by heightened clarity, strategic agency, and dialogic interaction. Thematic analysis initially yielded several recurring themes, which were subsequently organized into three overarching thematic clusters that reflect key areas of cognitive engagement: (1) clarity and accessibility, (2) strategic engagement and revision ownership, and (3) dialogic scaffolding through interactive feedback.

4.4.1. Clarity and Accessibility as a Catalyst for Engagement

A central difference between the two groups lay in the clarity and cognitive accessibility of feedback. Students in the ChatGPT-supported group frequently described the feedback as explicit, structured, and actionable. Student 4 (ChatGPT) noted, “It was straightforward—no guessing what it meant. I knew where to start revising.” Such clarity allowed learners to efficiently process and internalize suggestions, reducing the cognitive load typically associated with vague prompts.

In contrast, students in the teacher-supported group reported challenges interpreting feedback. Student 11 (Teacher) explained, “Sometimes it said ‘elaborate your ideas,’ but I didn’t know which part to elaborate or how.” Student 2 (Teacher) similarly reflected, “I understood that something was missing, but I couldn’t figure out what the teacher expected.” These observations reflect broader concerns in feedback literature about how ambiguous comments hinder learners’ ability to respond effectively [8,28]. Importantly, clarity was closely linked to perceived agency. As Student 5 (ChatGPT) observed, “When I could understand the feedback clearly, I felt I could own the revision process.” Thus, clarity did not merely enhance comprehension—it served as a gateway to revision autonomy.

4.4.2. Strategic Engagement and Ownership of Revision

Beyond understanding feedback, ChatGPT-supported students demonstrated more strategic and agentive revision behaviors. Rather than focusing solely on correcting surface errors, many described actively reorganizing content, refining logic, and improving rhetorical structure. Student 7 (ChatGPT) shared, “At first, I just fixed small grammar problems. But later, I looked at how to rearrange ideas or connect them better.” Similarly, Student 3 (ChatGPT) explained, “Before, I never thought about the order of my paragraphs. Now, I try to make the structure clearer for readers.”

These reflections parallel findings in the quantitative data showing a higher frequency of macrostructural changes among ChatGPT users. In contrast, teacher-supported students more often reported surface-level changes. Student 12 (Teacher) stated, “I still prefer just correcting words or grammar—it’s more straightforward.” Student 8 (Teacher) added, “I mostly followed what the teacher pointed out. I didn’t try to revise things they didn’t mention.”

Strategic revision was further reinforced by a growing sense of ownership and persistence. Student 9 (ChatGPT) described iterative revision as self-motivated: “Each time I saw something new I could improve.” Student 6 (ChatGPT) noted, “I didn’t just follow the advice—I tried different ways to express ideas.” These instances reflect cognitive agency, and align with [10] theory that mastery experiences foster self-efficacy. Student 5 (ChatGPT) summarized, “Now I feel like I have more control when I write. I can plan and revise without feeling lost.”

4.4.3. Dialogic Scaffolding Through Interactive Feedback

The third thematic cluster highlights the unique scaffolding potential of ChatGPT’s interactive format. Students perceived the feedback not as final evaluation, but as a collaborative dialogue. Student 14 (ChatGPT) described, “The way it explained what’s missing in my argument gave me a structure to work with.” Others noted that the multi-turn format enabled progressive refinement. As Student 10 (ChatGPT) recalled, “It was like a conversation—I could try something, and then get another suggestion.”

This iterative, responsive feedback allowed students to operate within their zone of proximal development [11], enabling revision steps beyond their independent capabilities. In contrast, teacher-supported students often experienced feedback as static and limited. Student 6 (Teacher) remarked, “Sometimes I knew what the problem was, but I didn’t know how to fix it.” Student 15 (Teacher) added, “I didn’t feel like I could ask again or clarify—it was just one-time advice.”

ChatGPT's nonjudgmental tone also reduced affective barriers. Student 7 (ChatGPT) observed, "It felt less like someone criticizing me, and more like someone guiding me." This emotional safety appeared to support sustained engagement, transforming revision into a process of co-construction rather than correction.

Together, these findings suggest that ChatGPT-supported feedback not only increased revision productivity but also enabled deeper cognitive investment in the writing process—supporting the multidimensional model of feedback engagement and reinforcing learners' self-efficacy development. By contrast, teacher-supported feedback, though appreciated for its authority, was often limited by interpretive ambiguity and a lack of interaction, which constrained students' ability to engage with revision tasks strategically and autonomously.

5. Discussion

This study set out to investigate how ChatGPT-supported feedback influences EFL students' writing development through a multi-dimensional lens, encompassing revision behavior, cognitive engagement, and writing self-efficacy. Situated within sociocognitive and sociocultural frameworks, the findings offer new insights into how generative AI tools may shift the role of feedback from static evaluation to dynamic interaction. These theoretical premises are increasingly reflected in empirical studies that highlight the potential of ChatGPT to serve as a scaffold for metacognitive engagement and learner-directed feedback use [4,29].

Findings addressing the first research question revealed that students who received ChatGPT-supported feedback produced significantly more revisions across all three writing tasks. More importantly, this productivity was sustained over time, signaling not mere compliance but the activation of a revision habit. In this regard, the feedback's immediacy and low-stakes format likely contributed to learners' willingness to experiment and persist in revision [30]. Emerging evidence supports the idea that AI-mediated feedback facilitates writing self-regulation and learner agency [25].

Building on this, the second research question showed that ChatGPT not only enhanced revision frequency, but also transformed revision quality. Students in the experimental group demonstrated significantly more macro- and micro-level changes, indicative of deeper rhetorical restructuring and ideational elaboration. Given that both groups started with equivalent pretest profiles, the observed differences underscore the specific affordances of AI-supported feedback in promoting cognitively complex revisions.

The behavioral data were mirrored in self-efficacy outcomes. Students in the ChatGPT group exhibited significantly greater increases across all three dimensions of writing self-efficacy—substantive, structural, and conventional—with large effect sizes. Self-efficacy tends to strengthen when learners experience mastery under supportive conditions [10]. The conversational and transparent feedback environment provided by ChatGPT likely enhanced students' sense of competence and control, thereby promoting confidence in planning, monitoring, and revising their work. In contrast, the teacher group's smaller gains—especially in discourse synthesis—suggest that feedback lacking individualized scaffolding may be less effective in building deeper writing confidence. ChatGPT's affordances—particularly its consistency, prompt elaboration, and nonjudgmental tone—may create a low-stakes, high-feedback environment that sustains learners' motivation and reduces cognitive inhibition, a key factor often overlooked in traditional writing classrooms.

Interview data further illuminated the mechanisms underlying these differences. Thematic analysis yielded three recurring clusters: (1) clarity and accessibility of feedback, (2) strategic engagement and revision ownership, and (3) dialogic scaffolding. Students in the ChatGPT group consistently emphasized the intelligibility and actionability of AI feedback, which reduced ambiguity and helped initiate revision. These findings resonate with the notion of feedback literacy, which highlights learners' need to interpret and use feedback independently [31]. Moreover, students frequently described experimenting with multiple versions—an indicator of cognitive agency and engagement. AI-enhanced environments that provide responsive and affectively neutral feedback appear to empower learners to act strategically, take ownership of revisions, and mitigate the anxiety often associated with teacher evaluations.

What most distinctly separated the two conditions was interactivity. ChatGPT's feedback unfolded as a dialogue, enabling learners to test ideas, receive elaborated suggestions, and build toward improved drafts. This iterative, scaffolded interaction mirrors the concept of mediated learning within the zone of proximal development [32]. Several students explicitly described ChatGPT as "less judgmental" and "more like a partner," emphasizing the role of emotional safety in fostering engagement. This aligns with a growing body of research that views AI tools as affective and cognitive facilitators in writing contexts [33].

Taken together, the findings suggest that ChatGPT functions not merely as a feedback provider, but as a dialogic co-participant. Its affordances—clarity, interactivity, and responsiveness—created a feedback ecology conducive to ownership, agency, and mastery. While teacher feedback was valued for its authority, its static and unilateral format often limited strategic engagement, particularly when it lacked elaboration or iterative interaction. This contrast underscores a broader pedagogical imperative to move beyond feedback quantity toward a rethinking of feedback quality and its interactional potential.

Pedagogically, these findings argue for the integration of generative AI tools into EFL writing instruction—not as a replacement for human feedback, but as a cognitive scaffold that complements teacher guidance. A hybrid instructional model may prove most effective: AI tools like ChatGPT can support lower-stakes, revision-focused feedback cycles, while teachers can prioritize higher-order concerns such as argumentation, genre awareness, and rhetorical sophistication [34]. However, such integration assumes a level of teacher readiness that cannot be taken for granted. As Patil et al. argue [35], targeted teacher training is essential to equip educators with the AI literacy and pedagogical competence needed to meaningfully incorporate generative AI tools into their instructional repertoire. Adaptive instructional models, where AI serves to mediate the revision process and educators engage in individualized conferencing, hold considerable promise. Future research might explore learner differences in AI uptake, longitudinal impacts on genre competence, and ethical concerns around overreliance. Importantly, task design should aim to cultivate feedback literacy—encouraging learners to critically negotiate, question, and adapt AI-generated feedback to foster reflective authorship rather than passive uptake [36].

6. Conclusions

This study examined the impact of ChatGPT-supported feedback on EFL learners' revision productivity, revision patterns, cognitive feedback engagement, and writing self-efficacy. Findings revealed that the ChatGPT-supported group outperformed the teacher-supported group in both the quantity and quality of revisions. Students in the ChatGPT group produced significantly more revisions and engaged in more content-level and macrostructural revisions, suggesting that AI-generated feedback may better facilitate knowledge-transforming writing. These effects were not simply behavioral; learners also reported higher levels of self-efficacy in planning, organizing, and refining their writing. Quantitative analysis showed significantly greater gains in all three dimensions of writing self-efficacy among ChatGPT users, with large effect sizes. These results align with emerging literature suggesting that AI-mediated feedback increases student confidence and engagement through iterative interaction [37]. Such tools, when designed to be transparent and responsive, can foster agency and autonomy by making feedback more interpretable and actionable.

Qualitative data further illuminated the mechanisms behind these outcomes. Thematic analysis of interviews revealed that ChatGPT feedback promoted deeper cognitive engagement through three interconnected processes: clarity and accessibility, strategic engagement and ownership, and dialogic scaffolding. Students described ChatGPT as not only understandable but also responsive and nonjudgmental—features that supported agency, experimentation, and emotional security. These findings resonate with recent research highlighting the affective benefits of dialogic AI interaction, particularly in promoting emotional safety during revision [38]. These characteristics helped scaffold learning by making expectations explicit and encouraging independent action. When AI feedback is intelligible and contextualized, it has the potential to enhance students' feedback literacy and empower them to revise with purpose and confidence [39]. In contrast, teacher feedback, though authoritative, was perceived as limited in elaboration and interactivity, sometimes hindering deeper revision thinking.

Taken together, these findings underscore the pedagogical potential of generative AI in writing instruction. Rather than serving as a replacement for teachers, tools such as ChatGPT may function as cognitive partners that enhance feedback uptake, foster emotional safety, and support learners' autonomy. Nonetheless, it is important to acknowledge that ChatGPT is only one instantiation of generative AI. Other platforms, such as DeepSeek or KIMI, may differ in interactive capacities, thereby limiting the generalizability of the present findings.

Future research should explore how learners with different proficiency levels, learning goals, or feedback preferences respond to AI-mediated revision support. Longitudinal studies could also investigate how repeated engagement with dialogic feedback affects writing development over time. Finally, further examination is needed into how students develop critical awareness and feedback literacy when interacting with AI tools.

Author Contributions

All authors have made a substantial, direct, and intellectual contribution to the work, including but not limited to Conceptualization, Methodology, Investigation, Formal analysis, Writing—Original Draft, and Writing—Review & Editing. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

All participants provided informed consent before participating in the study. The anonymity and confidentiality of the participants were guaranteed, and participation was completely voluntary.

Data Availability Statement

Data will be made available on request.

Acknowledgments

The authors appreciate the editors and all reviewers for their comments and suggestions.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix A

Interview Protocol

Section 1: Introduction

Welcome, state the interview's purpose and confirm confidentiality and consent;

Section 2: Questions

1. How do you usually understand and respond to feedback during writing?

2. Did the feedback help you clearly identify problems in your writing? Can you give an example?

3. When facing complex issues like paragraph structure or argument flow, was the feedback specific enough? How did you handle it?

4. What types of feedback most motivated you to revise? Why?

5. Did the feedback help you develop clearer ideas or writing structure? Please give an example.

6. Have you ever felt confused because the feedback was too vague? How did that affect your revision?

7. During revision, did you feel like you were "interacting" with the feedback? Was that process natural for you?

8. Did your revision behavior change over the course of the writing tasks? What do you think caused the change?

9. Looking back on the writing experience, have your writing confidence or ability changed? What role did feedback play in that?

Section 3: Closing

Thank the participant, confirm no further information, and reassure confidentiality.

References

[1] Li, A.W., Hebert, M., 2024. Unpacking an Online Peer-Mediated and Self-Reflective Revision Process in Second-Language Persuasive Writing. Reading and Writing. 37(6), 1545–1573. DOI: https://doi.org/10.1007/s11145-023-10466-8

[2] Tran, T.T.T., 2025. Enhancing EFL Writing Revision Practices: The Impact of AI- and Teacher-Generated

Feedback and Their Sequences. Education Sciences. 15(2), 232. DOI: https://doi.org/10.3390/educsci15020232

[3] Flower, L., Hayes, J.R., 1981. A Cognitive Process Theory of Writing. College Composition and Communication. 32(4), 365–387. DOI: https://doi.org/10.58680/ccc198115885

[4] Guo, K., Zhang, E.D., Li, D., et al., 2024. Using AI-Supported Peer Review to Enhance Feedback Literacy: An Investigation of Students' Revision of Feedback on Peers' Essays. British Journal of Educational Technology. 6(4), 1612–1639. DOI: https://doi.org/10.1111/bjet.13540

[5] Hyland, K., Hyland, F., 2006. Feedback on Second Language Students' Writing. Language Teaching. 39(2), 83–101. DOI: https://doi.org/10.1017/S0261444806003399

[6] Han, Y., Hyland, F., 2015. Exploring Learner Engagement With Written Corrective Feedback in a Chinese Tertiary EFL Classroom. Journal of Second Language Writing. 30(1), 31–44. DOI: https://doi.org/10.1016/j.jslw.2015.08.002

[7] Zhu, Q., Carless, D., 2018. Dialogue Within Peer Feedback Processes: Clarification and Negotiation of Meaning. Higher Education Research and Development. 37(4), 883–897. DOI: https://doi.org/10.1080/07294360.2018.1446417

[8] Carless, D., Boud, D., 2018. The Development of Student Feedback Literacy: Enabling Uptake of Feedback. Assessment and Evaluation in Higher Education. 43(8), 1315–1325. DOI: https://doi.org/10.1080/02602938.2018.1463354

[9] 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. DOI: https://doi.org/10.29140/dal.v2.102469

[10] Bandura, A., 1997. Self-Efficacy: The Exercise of Control. W.H. Freeman: New York, NY, USA. pp. 1–604.

[11] Vygotsky, L.S., 1978. Mind in Society: The Development of Higher Psychological Processes. Harvard University Press: Cambridge, NY, USA. pp. 1–159.

[12] Tajik, A., 2025. Integrating AI-Driven Emotional Intelligence in Language Learning Platforms to Improve English Speaking Skills Through Real-Time Adaptive Feedback. PREPRINT (Version 1). DOI: https://doi.org/10.21203/rs.3.rs-5919944/v1(cited 1 Match 2025).

[13] Ferris, D.R., 2006. Does Error Feedback Help Student Writers? New Evidence on Short- and Long-Term Effects of Written Error Correction. In: Hyland, K., Hyland, F. (eds.). Feedback in Second Language Writing: Contexts and Issues. Cambridge University Press: Cambridge, UK. pp. 81–104.

[14] Ferris, D.R., 2014. Responding to Student Writing: Teachers' Philosophies and Practices. Assessing Writing. 19, 6–23. DOI: https://doi.org/10.1016/j.asw.2013.09.004

[15] Wu, Y., Schunn, C.D., 2020. From Feedback to Revisions: Effects of Feedback Features and Perceptions. Contemporary Educational Psychology. 60, 101826. DOI: https://doi.org/10.1016/j.cedpsych.2019.101826

[16] Yang, L., Li, R., 2024. ChatGPT for L2 Learning: Current Status and Implications. System. 124, 103351. DOI: https://doi.org/10.1016/j.system.2024.103351

[17] Liu, C., Yu, S., 2022. Reconceptualizing the Impact of Feedback in Second Language Writing: A Multidimensional Perspective. Assessing Writing. 52, 100630. DOI: https://doi.org/10.1016/j.asw.2022.100630

[18] Du, J., Alm, A., 2024. The Impact of ChatGPT on English for Academic Purposes (EAP) Students' Language Learning Experience: A Self-Determination Theory Perspective. Education Sciences. 14(7), 726. DOI: https://doi.org/10.3390/educsci14070726

[19] Guo, X., 2024. Facilitator or Thinking Inhibitor: Understanding the Role of ChatGPT-Generated Written Corrective Feedback in Language Learning. Interactive Learning Environments. 1–19. DOI: https://doi.org/10.1080/10494820.2024.2445177

[20] Liu, Z.M., Hwang, G.J., Chen, C.Q., et al., 2024. Integrating Large Language Models Into EFL Writing Instruction: Effects on Performance, Self-Regulated Learning Strategies, and Motivation. Computer Assisted Language Learning. 1–25. DOI: https://doi.org/10.1080/09588221.2024.2389923

[21] Banaganipalli, B.N.P., 2025. ChatGPT as a Scaffolding Tool in English Language Teaching. Literary Vibes. 14(1), 1–7. DOI: https://doi.org/10.5281/zenodo.14898233

[22] Creswell, J.W., Plano Clark, V.L., 2018. Designing and Conducting Mixed Methods Research, 3rd ed. SAGE Publications: Thousand Oaks, USA. pp. 1–492.

[23] Creswell, J.W., Poth, C.N., 2018. Qualitative Inquiry and Research Design: Choosing Among Five Approaches, 4th ed. SAGE Publications: Thousand Oaks, USA. pp. 1–488.

[24] Mackey, A., Gass, S.M., 2016. Second Language Research: Methodology and Design, 2nd ed. Routledge: New York, USA. DOI: https://doi.org/10.4324/9781315750606

[25] Zhan, Y., Yan, Z., 2025. Students' Engagement With ChatGPT Feedback: Implications for Student Feedback Literacy in the Context of Generative Artificial Intelligence. Assessment and Evaluation in Higher Education. Assessment & Evaluation in Higher Education. 1–14. DOI: https://doi.org/10.1080/02602938

.2025.2471821
[26] Field, A., 2018. Discovering Statistics Using IBM SPSS Statistics, 5th ed. SAGE Publications: London, UK. pp. 1–915.
[27] Braun, V., Clarke, V., 2006. Using Thematic Analysis in Psychology. Qualitative Research in Psychology. 3(2), 77–101. DOI: https://doi.org/10.1191/1478088706qp063oa
[28] Sweller, J., 2011. Cognitive load theory. Psychology of Learning and Motivation. 55, 37–76. DOI: https://doi.org/10.1016/B978-0-12-387691-1.00002-8
[29] Tajik, A., 2025. Beyond Voice Recognition: Integrating Alexa's Emotional Intelligence and ChatGPT's Language Processing for EFL Learners' Development and Anxiety Reduction—A Comparative Analysis. 11 February 2025, PREPRINT (Version 1). DOI: https://doi.org/10.21203/rs.3.rs-5989702/v1
[30] Faigley, L., Witte, S., 1981. Analyzing Revision. College Composition and Communication. 32(4), 400–414. DOI: https://doi.org/10.58680/ccc198115887
[31] Ketonen, L., 2021. Exploring Interconnections Between Student Peer Assessment, Feedback Literacy and Agency [Master's thesis]. University of Jyväskylä: Jyväskylä, Finland.
[32] Xie, X., Yong, M.F., Yap, N.T., 2024. Students' Perceptions of Evaluative Judgement in Technology-Mediated Dialogic Peer Feedback. Pertanika Journal of Social Sciences and Humanities. 32(4), 1643–1660. DOI: https://doi.org/10.47836/pjssh.32.4.19
[33] Rad, H.S., Mirzaei, A., 2024. Developing Feedback Literacy, Scaffolded Writing, and Resilience Through Intervention on Feedback Processes for L2 Writing Students. Studies in Educational Evaluation. 81, 101354. DOI: https://doi.org/10.1016/j.stueduc.2024.101354
[34] Choi, M.S., 2024. Enhancing Students' Writing Skills Through a Dialogic Process With Instructional AI in the Post-Pandemic Era. In: Tomei, L.A., Carbonara, D.D. (eds.). Exploring Technology-Infused Education in the Post-Pandemic Era. IGI Global: Hershey, USA. pp. 268–283. DOI: https://doi.org/10.4018/979-8-3693-2885-9.ch008
[35] Patil, N.G., Kou, N.L., Baptista-Hon, D.T., et al., 2025. Artificial intelligence in medical education: A practical guide for educators. MedComm-Future Medicine. 4(2), e70018. DOI: https://doi.org/10.1002/mef2.70018
[36] Wood, J., 2021. A Dialogic Technology-Mediated Model of Feedback Uptake and Literacy. Assessment and Evaluation in Higher Education. 46(8), 1173–1190. DOI: https://doi.org/10.1080/02602938.2020.1852174
[37] Daniel, K., Msambwa, M.M., Wen, Z., et al., 2025. Can Generative AI Revolutionise Academic Skills Development in Higher Education? A Systematic Literature Review. European Journal of Education. 60(1), e70036. DOI: https://doi.org/10.1111/ejed.70036
[38] Derakhshan, A., 2025. EFL Students' Perceptions About the Role of Generative Artificial Intelligence (GAI)-Mediated Instruction in Their Emotional Engagement and Goal Orientation: A Motivational Climate Theory (MCT) Perspective in Focus. Learning and Motivation. 90, 102114. DOI: https://doi.org/10.1016/j.lmot.2025.102114
[39] Agostini, D., Picasso, F., 2024. Large Language Models for Sustainable Assessment and Feedback in Higher Education: Towards a Pedagogical and Technological Framework. Intelligenza Artificiale. 18(1), 121–138. DOI: https://doi.org/10.3233/IA-240033