Journal of Education and Training Studies
Vol. 13, No. 4; October 2025
ISSN 2324-805X E-ISSN 2324-8068
Published by Redfame Publishing
URL: http://jets.redfame.com

Perceptions and Effectiveness of AI-Assisted Written Corrective Feedback: *A Case Study of Chinese EFL University Students*

Shangrong Li

Correspondence: Shangrong Li, School of International Languages and Cultures, Hohai University, Nanjing, China.

Received: May 5, 2025 Accepted: June 15, 2025 Online Published: June 22, 2025

doi:10.11114/jets.v13i4.7784 URL: https://doi.org/10.11114/jets.v13i4.7784

Abstract

This study explores Chinese university students' attitudes toward AI-assisted written corrective feedback (AI-WCF) and their perceived improvement in argumentative writing. Adopting a one-group retrospective pretest–posttest design, the study collected data from 89 first-year undergraduates through two questionnaires: one on attitudes toward AI-WCF, and the other on self-assessed progress across four writing dimensions—unity, support, cohesion and coherence, and language use. Over a 12-week College English course, students engaged in structured writing tasks supported by AI feedback. Results showed generally positive attitudes, with strong endorsement of AI-WCF's usefulness and future applicability. Paired-samples t-tests revealed significant perceived gains across all subskills ($p < .001$), with large effect sizes ($d = 0.77–1.57$), particularly in thematic clarity, organization, and lexical precision. Nonetheless, students expressed reservations about fully integrating AI-WCF into formal instruction, citing its limited rhetorical depth. Findings highlight the importance of differentiated support and suggest integrating AI tools within scaffolded feedback processes to enhance learner engagement and autonomy.

Keywords: written corrective feedback, AI-assisted, attitude, perceived effectiveness, L2 writing

1. Introduction

Written corrective feedback (WCF) has long been recognized as a key instructional strategy in second language (L2) writing classrooms, helping learners identify linguistic errors, refine rhetorical structures, and improve overall writing competence (Bitchener & Ferris, 2012; Hyland & Hyland, 2006). WCF has been traditionally delivered by teachers in the form of direct or indirect comment regarding grammar, vocabulary, coherence, and organizational aspects. In large classes, however, teacher feedback is often inconsistent, delayed, and lacks customization (Ferris, 2010), creating demand for more scalable and effective alternatives.

The fast evolution of artificial intelligence (AI), particularly with the advent of big language models (LLMs) such as ChatGPT, Notion AI, and Grammarly, has opened new avenues to providing automatic, immediate, and tailored writing feedback. AI-assisted written corrective feedback (AI-WCF) is automatically generated feedback presented by AI systems to assist learners in revising what they write in relation to grammatical correctness, word choice, sentence smoothness, content pertinence, and structural clarity (Ranalli, 2018). Because AI-WCF is more convenient, uniform, and immediate than teacher feedback, it may enhance learner autonomy and engagement with feedback (Tran, 2025; Osawa, 2023).

Recent research has yielded preliminary evidence of the educational value of AI-WCF. On the one hand, empirical evidence is indicative of how AI tools promote greater accuracy in writing, less anxiety, and greater self-confidence among L2 learners (Athanassopoulos et al., 2023; Wang, 2024). On the other hand, comparative studies have revealed that while teachers are capable of providing more in-depth rhetorical and content-level feedback, AI systems are better equipped to offer direct, metalinguistic correction on surface-level aspects like grammar and mechanics (Lin & Crosthwaite, 2024; Ren, Liu, & Xie, 2024). In blended feedback contexts, combining teacher input with AI suggestions has been found to enhance revision quality and increase student engagement, particularly in large classes (Wiboolyasarin et al., 2024; Jiang, 2025). Some studies also highlight the instructional benefits of integrating AI-WCF with tools like e-portfolios to support learner autonomy and feedback literacy (Osawa, 2023). Another emerging line of inquiry focuses on learner engagement with and reactions to AI feedback. Chen, Zhu, Lu and Wei (2024) found that students often accept surface-level feedback from ChatGPT but may reject or ignore suggestions related to content and argumentation due to perceived irrelevance or lack of specificity. These findings emphasize the need to explore not just the content of feedback but also learners' cognitive and emotional responses to AI-generated input.

Despite these advances, two important research gaps remain. First, most existing studies rely on product-based assessments such as writing scores or revision accuracy, while overlooking learners' subjective experiences—including satisfaction, emotional attitudes, and perceived learning gains. These affective attitudes and self-perceived skill improvements are distinct yet interrelated: affective attitudes influence a learner's engagement and openness to feedback, while perceived skill gains reflect the learner's judgment of their own development as a result of feedback. Both are crucial for understanding how feedback is processed and utilized (Gardner, 1985; Davis, 1989). Second, there is minimal understanding regarding how students assess the efficiency of AI-WCF with reference to overall qualities of academic writing like unity, support, coherence, and language use, which are central to argumentative writing and are in close relation to instructional goals.

To address these gaps, this study explores Chinese university learners' affective attitudes toward AI-supported written corrective feedback and their perceived improvement in academic writing. Employing a one-group retrospective pretest–posttest research design (Little, 2019; Li, 2024), the research investigates learners' self-reported reflections on their writing ability before and after exposure to AI feedback, as well as their reviews of the feedback process itself. Unlike prior studies that rely chiefly on pre- and post-intervention test scores, this study shifts the focus to student-perceived efficacy and the actual uptake of AI-generated feedback—examining not only how students feel about using AI-WCF but also how they perceive its impact on their writing skills. By examining both emotional responses and perceived skill improvements, this study provides new insights into how learners interact with and benefit from AI feedback in real classroom contexts and offers practical implications for instructional practice.

2. Literature Review

2.1 Written Corrective Feedback (WCF) in L2 Writing

Written corrective feedback (WCF) involves responses given to learners' written work in the hope of making learners notice, recognize, and amend grammatical, vocabulary, organizational, and content errors (Bitchener & Ferris, 2012). As one of the fundamental aspects of second language (L2) writing instruction, WCF functions in two ways: in supporting immediate text repair and in aiding long-term language acquisition. WCF is normally given in one of several forms—direct correction, error codes, metalinguistic explanations, or reformulations—and may address surface-level accuracy or more advanced writing considerations such as cohesion and rhetorical style (Ellis, 2009).

Despite extensive empirical evidence supporting the efficacy of WCF in enhancing writing accuracy and complexity (Ferris, 2010; Hyland & Hyland, 2006), its use in actual classrooms is frequently constrained by time pressures, teacher workload, and high student enrollments. These constraints have led researchers and teachers to investigate alternative methods of providing feedback, specifically through technology-facilitated means.

2.2 The Rise and Application of AI-assisted Written Corrective Feedback

The recent growth in artificial intelligence (AI), particularly in large language models like ChatGPT, Grammarly, and Notion AI, has created new opportunities for automating WCF. Automatically produced suggestions relating to errors in grammar, clarity, coherence, lexical selection, and text organization have been termed as AI-assisted written corrective feedback (AI-WCF) (Ranalli, 2023). These tools present scalable and immediate feedback, with the potential to enhance learner autonomy and motivation (Osawa, 2023;Wang, 2024).

Empirical studies support the pedagogical value of AI-WCF. Wang (2024) demonstrated that AI feedback reduces writing anxiety and enhances both fluency and accuracy. Athanassopoulos et al. (2023) similarly found that ChatGPT-supported revision led to vocabulary and grammar improvement among multilingual learners. Lin and Crosthwaite (2024), comparing teacher and AI feedback, noted that while teacher comments were richer in content-level suggestions, AI feedback tended to be more consistent and focused on form. Hybrid feedback models have also gained traction. Han and Li (2024) observed improved learner uptake when teacher and AI feedback were combined. Jiang (2025), using the ICAP framework, reported that integrated teacher–AI feedback enhanced interaction and writing outcomes. In parallel, Osawa (2023) explored the use of Notion AI within e-portfolio systems, finding that such integration supported self-regulated learning and scaffolded reflection.

2.3 Learners' Attitudes toward AI-assisted Written Corrective Feedback

Learner attitudes play a crucial role in the success of any instructional intervention, particularly those involving new technologies. In Gardner's (1985) model of socio-education and in Davis's (1989) technology acceptance model (TAM), learners' intentions and motivation are largely determined by how useful, easy to use, and relevant to themselves the technology is perceived to be. L2 writing is a context where these perceptions are highly appropriate as learners are introduced to AI-supported written corrective feedback (AI-WCF) tools, which will presumably be very different in delivery manner and focus from traditional teacher corrective feedback.

More recent studies have started to investigate learners' affective and evaluative responses to AI feedback. Wang (2024) found that, overall, learners perceived AI feedback as helpful and not threatening, specifically in minimizing writing apprehension. Lin and Crosthwaite (2024) determined that learners valued the immediacy and directness of ChatGPT feedback but were concerned with its redundancy and inability to address deeper-level issues with content. Osawa (2023) further highlighted that learners were receptive to the unobtrusive embedding of Notion AI in writing environments such as e-portfolios, especially with self-regulation scaffolds. Chen et al. (2024) further established that learners also resisted AI feedback in cases where it lacked clarity or seemed to misrepresent their intended meaning—indicating that learner attitudes toward AI feedback may be complex and context-specific. Yet most of these studies assess attitudes indirectly or as secondary measures, in many cases in performance-based research studies. There are few systematic studies explicitly addressing how student learners assess the overall experience with AI-WCF and whether and how likely they are to adopt it in the future.

Since student attitudes not only affect participation but also extent of usage, greater understanding of student perceptions is necessary in order to successfully implement AI feedback in writing instruction.

2.4 Perceived Effectiveness of AI-Aided Written Corrective Feedback

Whereas learner attitudes capture how students feel about AI-supported feedback, perceived effectiveness refers to how they assess its impact on their writing development. Existing studies have increasingly examined the effectiveness of AI-WCF using a range of research designs. For example, Lin and Crosthwaite (2024) found that ChatGPT's feedback, while metalinguistic and clear, could be redundant and sometimes inconsistent. Wang (2024) demonstrated that AI-generated feedback significantly enhanced learners' fluency and accuracy and reduced apprehension. Chen et al. (2024) found students more likely to reject AI feedback when it addressed content, citing ambiguity and lack of pertinence. Godwin-Jones (2024) highlighted the need for teacher mediation to maximize AI's benefits for learners. These findings all support the promise of AI-WCF to facilitate enhancements in language accuracy, coherence, and fluency—particularly with AI tools employed on form-focussed revisions. They also, though, highlight shortfalls, such as superficiality in addressing concerns of higher order in writing and learner reluctance to accept AI feedback on idea conception or rhetorical makeup.

In spite of these efforts, most studies utilize performance-based measures (e.g., pre-post-test scores, quality of revisions) as proxies for effectiveness. Although these methods are objective, they may not reflect learners' own self-perceptions or the subtle ways in which AI responses are incorporated and utilized. Additionally, standard pretest–posttest designs are vulnerable to response-shift bias, where learners' conception of target competencies shifts with time—possibly distorting the accuracy of early ratings and compromising the validity of longitudinal comparisons (Little et al., 2019; Sprangers & Schwartz, 1999).

In order to bridge these gaps, the current research uses the retrospective pretest–posttest (RPP) model with an attitude survey, mirroring the learner-centered approach advocated by Little et al. (2019). This is possible because it permits valid and reliable assessment of perceived learning gains without disrupting instruction flow. Targeting particular subskills—namely, unity, support, cohesion and coherence, and language use—that are most relevant to course instruction objectives, this research provides a more refined analysis of AI-WCF's perceived effectiveness as part of academic writing improvement.

Thus, the research seeks to tackle the following research questions:
(1) What are the students' attitude toward AI-assisted WCF in academic writing?
(2) How do students perceive the effect of AI-assisted WCF on their writing performance?

3. Method

3.1 Research Design

A quantitative survey design using a one-group retrospective pretest-posttest (RPP) design was adopted, exploring students' attitudes toward AI-assisted WCF (AI-WCF) and their perceptions of academic writing improvement. We chose the RPP model rather than traditional pretest–posttest (TPP) models due to its strengths in assessing subjective constructs, such as perceptions, attitudes, and self-report of competence (Bray, Maxwell, & Howard, 1984; Howard, Dailey, & Gengler, 1979).

In contrast to the TPP designs, where the along-group pretest can lead to participants' irrational overestimation of their initial abilities due to naïveté, the RPP model instructs participants to make retrospective assessment of their posttest performance at the same time point their posttest is administered, that is, after the learning experience. This minimizes response-shift bias and provides more valid measurement of perceived change (Sprangers & Schwartz, 1999). As an approach to classroom-based educational research, the RPP model has also been found to be cost-effective and pedagogically non-invasive (Little et al., 2019), specifically in L2 settings where metacognitive reflection is vital for learner advancement (Li, 2024).

To complement the self-reported gains via the RPP instrument, a one-time attitude survey was conducted to capture the evaluative reactions of students to AI-WCF (i.e., perceived usefulness, satisfaction, and intention with regard to future use). With these two instruments integrated, the design of the study can offer a better perspective regarding the ways in which learners experienced, internalized, and evaluated the AI-generated feedback in a real learning environment.

3.2 Context and Participants

This study was conducted at a university in East China, involving 89 first-year undergraduate students majoring in non-English disciplines. All participants were enrolled in a comprehensive College English course during their first semester. The course was delivered over 12 weeks, comprising a total of 48 class hours (four hours per week), and integrated instruction in listening, speaking, reading, and writing. The writing component of the course focused on argumentative writing and was structured around four key units: unity, support, cohesion and coherence, and language use. Each unit spanned two weeks. In class, the instructor introduced writing strategies relevant to the unit focus and provided scaffolded prompt templates—instructor-designed model queries or sentence stems based on each unit's writing objectives, aimed at helping students interact productively with AI. For instance, for the cohesion unit, a prompt might be: "Please use cohesive devices, such as reference, substitution, ellipsis, or conjunctions, to make the connections between my sentences and paragraphs clearer." These templates were used to guide students in planning and improving their texts. Following class time, students did a writing task and employed AI tools to solicit feedback. Students were asked to engage with AI based upon the given prompts, correct their drafts following the AI feedback, and submit their end versions for teacher comments and grading. For Weeks 1–2, instruction on how to engage with AI tools was given to the students. For Weeks 3–10, students worked on writing assignments corresponding to each unit, corrected their drafts according to AI feedback, and got teacher comments and grading for the end versions. For Week 11, both attitude questionnaire and retrospective pretest–posttest questionnaire were given to examine students' attitudes and their self-assessed improvement.

Among the 89 participants, 65were male (73%) and 24 were female (27%). None had overseas study experience. All students had received over ten years of systematic English instruction under China's national curriculum, with an emphasis on reading and writing. Their English proficiency was measured by the College English Test Band 4 (CET-4), with an average score of 485.4 (SD = 61.1), indicating an intermediate level of competence.

3.3 Instruments

Two questionnaires were designed to gather data on students' experiences using AI-assisted corrective written feedback (AI-WCF) and students' perception of improvement in their academic writing. The two instruments used a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). The questionnaires were developed based on the course instructional goals and literature on learner self-assessment and feedback perception.

3.3.1 Attitude Questionnaire

The attitude survey was employed to measure students' overall experience of AI-based written corrective feedback. It included four items that gauged learners' affective reactions, usefulness, value within the curriculum, and intention to use AI tools for future writing assignments. Sample items are: "My experience of AI-assisted written corrective feedback was positive" and "I would like to continue to use AI-assisted written corrective feedback in my future writing." The Cronbach's alpha was 0.76 , indicating a satisfactory internal consistency of the attitude scale.

3.3.2 Perceived Effectiveness Questionnaire

A retrospective pretest–post-test survey was constructed for examining students' perceptions of the effectiveness of AI-assisted WCF to improve writing skills. This instrument was developed based on the rationale for the reflective self-assessment task devised by Li (2024) as it has been found especially appropriate in capturing the perceived learning gains in classroom-based settings where conventional pretesting may disrupt instruction or result in inauthentic self-ratings. In this survey, subjective improvement in two core modules—skills and knowledge—was assessed. The skills module was organized around the four fundamental elements of argumentative writing that were emphasized in the course: unity, support, cohesion and coherence, and language use. At each item, learners were asked to report their level of proficiency (retrospective pretest and posttest) using AI intervention. For example, within the "unity" area, learners measured items like "I have a clearly stated thesis in the introduction." The knowledge module featured items for students' recognition of composition strategies consistent with every unit target. The survey evidenced high internal consistency across both the retrospective pretest ($\alpha = 0.89$) and posttest ($\alpha = 0.90$).

3.4 Data Collection and Analysis

Data were collected during Week 11 of the semester through two online questionnaires administered via Tencent Questionnaire, a secure digital survey platform. Participation in the study was voluntary and anonymous, and informed consent was obtained from all students prior to data collection. The attitude questionnaire was administered once to evaluate students' overall experience with AI-assisted written corrective feedback (AI-WCF). The perceived effectiveness questionnaire employed a retrospective pretest–posttest format, requiring participants to assess their writing knowledge and skills both "then" (before receiving AI feedback) and "now" (after the intervention), based on their reflective judgment at the end of the course.

Descriptive statistics, including means and standard deviations, were calculated to summarize students' attitudes and self-assessed improvement. Paired-samples t-tests were used to examine statistically significant differences between the retrospective pretest and posttest ratings across specific writing knowledge and writing sub-skills. Cronbach's alpha coefficients were computed to evaluate the internal consistency of both instruments. All quantitative analyses were conducted using SPSS (version 26).

4. Results and Discussion

4.1 Students' Attitudes toward AI-assisted WCF

Descriptive statistics from the attitude questionnaire revealed that students generally held positive views toward AI-assisted written corrective feedback (AI-WCF). As shown in Table 1, all four items received mean scores above the midpoint of the 5-point Likert scale (3.00), indicating favorable evaluations across dimensions of satisfaction, perceived usefulness, curricular value, and future intent.

Table 1. Descriptive statistics of students' attitudes toward AI-assisted written corrective feedback (n = 89)

Items	Min	Max	Mean	SD
My experience of AI-assisted written corrective feedback was positive.	2.0	5.0	3.944	0.663
I think AI-assisted written corrective feedback was helpful to improve my argumentative writing skills.	2.0	5.0	4.011	0.666
I think it was worthwhile to integrate AI-assisted written corrective feedback into our course.	1.0	5.0	3.764	0.84
I would like to continue to use AI-assisted written corrective feedback in my future writing.	2.0	5.0	4.079	0.742

The highest mean score was observed for the item "I would like to continue to use AI-assisted feedback in my future writing" (M = 4.08, SD = 0.74), followed closely by "I think AI-assisted written corrective feedback was helpful to improve my argumentative writing skills" (M = 4.01, SD = 0.67). These results suggest strong learner recognition of AI-WCF's value in supporting their writing development and a willingness to continue using such tools beyond the current course. The item "My experience of AI-assisted written corrective feedback was positive" also received a high average score (M = 3.94, SD = 0.66), indicating a generally positive emotional response. However, the lowest mean score was for "I think it was worthwhile to integrate AI-assisted written corrective feedback into our course" (M = 3.76, SD = 0.84), showing that while students are personally receptive to AI-WCF, they may hold a more cautious attitude toward its formal curricular integration.

These results align with prior findings on student receptivity to AI-assisted feedback. For instance, Wang (2024) similarly reported high student satisfaction and reduced writing anxiety when using ChatGPT-generated feedback, attributing this to the non-threatening, consistent, and immediate nature of AI interaction. Lin and Crosthwaite (2024) also found that learners valued the clarity and speed of ChatGPT's suggestions, even while expressing skepticism about its depth and relevance for complex rhetorical revisions. The strong endorsement of continued use in future writing (M = 4.08) echoes findings from Osawa (2023), where students voluntarily integrated Notion AI into e-portfolio projects after initial exposure. This suggests that once learners are introduced to AI tools in a guided context, they tend to appreciate the autonomy and flexibility afforded by such systems and therefore would like to keep using them in the future. The relatively lower score for curricular integration (M = 3.76), however, points to ongoing reservations. This cautious stance may stem from limitations in AI feedback's handling of global-level writing concerns—such as idea development, argument logic, or discipline-specific conventions—as previously reported by Chen et al. (2024) and Ren et al. (2024). Students may recognize that while AI is useful for surface-level revisions and structural clarity, it lacks the pedagogical depth, adaptability, and contextual sensitivity offered by human instructors. These mixed attitudes imply that AI-WCF should be positioned not as a replacement but as a complementary tool to teacher feedback. Its integration may be most effective when embedded into a scaffolded writing process where students are taught how to interpret, evaluate, and selectively apply AI-generated suggestions. Such an approach could not only address learners' practical needs but also build their feedback literacy and critical thinking skills.

Future instructional interventions could consider offering tiered feedback models—for example, using AI for initial drafting and revision, followed by teacher feedback for content, logic, and discourse. Additionally, ongoing teacher support in prompt engineering and AI evaluation strategies may help mitigate learner skepticism and enhance instructional trust in AI systems.

4.2 Perceived Effectiveness of AI-assisted WCF on Writing Knowledge and Skills

The results of paired-samples t-tests (as shown in Table 2) revealed statistically significant improvements across all items in the retrospective pretest–posttest questionnaire. Students perceived notable gains in both writing knowledge and skills after engaging with AI-assisted written corrective feedback (AI-WCF). All comparisons yielded $p < .001$, with large effect sizes according to Cohen's d, ranging from 0.77 to 1.57. These results indicate not only statistical significance but also strong educational relevance.

Table 2. Paired-samples t-test results for students' perceived progress in writing knowledge and skills (n = 89)

Items	Now M (SD)	Then M (SD)	t	p	d
Know strategies for maintaining thematic unity (e.g., thesis and topic sentences).	3.91 (0.65)	2.71 (0.87)	-15.30	< .001	1.57
Identify effective types of supporting evidence (e.g., examples, statistics, expert testimony).	3.51 (0.84)	2.39 (0.98)	-12.04	< .001	1.22
Understand cohesion and coherence-building techniques (e.g., transitions, logical organization).	3.88 (0.70)	2.84 (0.88)	-11.36	< .001	1.30
Know strategies for lexical and syntactic refinement (e.g., specific words, varied sentence structures).	3.27 (0.84)	2.30 (0.80)	-10.58	< .001	1.21
Construct a clear thesis in the introduction.	3.88 (0.84)	2.87 (1.10)	-11.46	< .001	1.06
Write topic sentences for each paragraph.	3.58 (0.86)	2.51 (1.08)	-10.64	< .001	1.08
Delete irrelevant content that does not support the thesis.	3.24 (0.93)	2.30 (0.99)	-9.86	< .001	1.04
Provide specific and persuasive supporting details.	3.21 (0.91)	2.18 (0.87)	-11.64	< .001	1.20
Use appropriate transition words to improve flow.	3.72 (0.72)	2.79 (0.96)	-11.27	< .001	1.10
Apply logical sequencing for coherence (e.g., comparison, cause-effect).	3.27 (0.89)	2.36 (0.83)	-10.29	< .001	1.07
Use specific rather than vague vocabulary.	3.20 (0.83)	2.21 (0.83)	-10.86	< .001	1.19
Avoid wordiness and use concise expressions.	3.28 (0.81)	2.40 (0.90)	-10.52	< .001	1.09
Use varied sentence structures.	3.39 (0.82)	2.40 (0.93)	-13.05	< .001	1.25
Check spelling, punctuation, and capitalization.	3.75 (0.90)	3.02 (1.06)	-8.38	< .001	0.77

In the knowledge domain, the greatest improvement was found in students' awareness of how to maintain thematic focus, which increased from M = 2.71 (SD = 0.87) to M = 3.91 (SD = 0.65), $t(88) = -15.30$, $p < .001$, d = 1.57. Similarly, knowledge of effective types of supporting evidence improved from M = 2.39 (SD = 0.98) to M = 3.51 (SD = 0.84), $t(88) = -12.04$, $p < .001$, d = 1.22. Students also reported better understanding of cohesion and coherence-building strategies such as logical sequencing and use of transitions, with scores rising from M = 2.84 (SD = 0.88) to M = 3.88 (SD = 0.70), $t(88) = -11.36$, $p < .001$, d = 1.30. These findings suggest that AI tools provided clear and accessible suggestions that supported students' rhetorical development, particularly at the organizational and structural levels of writing.

In terms of writing skills, students perceived substantial growth in their ability to construct clear thesis statements, with scores rising from M = 2.87 (SD = 1.10) to M = 3.88 (SD = 0.84), $t(88) = -11.46$, $p < .001$, d = 1.06. Similarly, topic sentence generation improved from M = 2.51 (SD = 1.08) to M = 3.58 (SD = 0.86), $t(88) = -10.64$, $p < .001$, d = 1.08. Notable progress was also observed in cohesion and coherence. For instance, the ability to use transitions improved from M = 2.79 (SD = 0.96) to M = 3.72 (SD = 0.72), $t(88) = -11.27$, $p < .001$, d = 1.10, and logical sequencing from M = 2.36 (SD = 0.83) to M = 3.27 (SD = 0.89), $t(88) = -10.29$, $p < .001$, d = 1.07. Language-focused sub-skills also demonstrated strong gains. Lexical specificity improved from M = 2.21 (SD = 0.83) to M = 3.20 (SD = 0.83), $t(88) = -10.86$, $p < .001$, d = 1.19; sentence variety from M = 2.40 (SD = 0.93) to M = 3.39 (SD = 0.82), $t(88) = -13.05$, $p < .001$, d = 1.25. Among all items, the smallest gain was observed in language mechanics (e.g., spelling, punctuation), where scores rose from M = 3.02 (SD = 1.06) to M = 3.75 (SD = 0.90), $t(88) = -8.38$, $p < .001$, d = 0.77.

These findings confirm that AI-assisted WCF led to substantial learner-perceived improvements across both declarative writing knowledge and practical writing skills. The large effect sizes observed for most items (d > 1.0) underscore not only statistically significant changes but also meaningful pedagogical impact. This is in line with previous studies (e.g., Wang, 2024; Wei & Li, 2023), which have shown that AI feedback promotes greater learner autonomy and facilitates revision in academic writing. The most notable perceived gains occurred in higher-order organizational and rhetorical aspects—such as thesis development, paragraph structuring, and logical flow. These findings resonate with Ren et al. (2024), who reported that ChatGPT tends to provide consistent, structured feedback on global-level writing concerns.

AI's ability to scaffold writing tasks through directive prompts likely contributed to students' improved awareness of argument structure. Improvements in surface-level language dimensions, such as lexical precision and sentence variation, were also robust, supporting Lin and Crosthwaite's (2024) observation that AI systems are particularly effective in identifying fluency-level issues and suggesting local edits. However, the relatively smaller effect size for mechanics (d = 0.77) may reflect a ceiling effect or limited learner attention to lower-order issues during AI-mediated revision.

Importantly, although overall outcomes were positive, variation in standard deviations and minimum scores indicates considerable individual differences in perceived improvement. This reinforces the idea that AI-WCF, while broadly effective, may not equally benefit all students. Such disparities may stem from differences in baseline proficiency, feedback processing strategies, or engagement levels. Therefore, differentiated AI-mediated support from the teachers should be incorporated to classroom instruction. First, instructors can diagnose both the actual and potential writing competence of students through pre-assessment tasks and reflective self-checks. This diagnosis can, in turn, assist the teachers in understanding in which situation the students are at the present moment and how to scaffold their progress. Based on this assessment, the teachers can generate tailoring prompts for the students according to their specific writing problems. For instance, a student struggling with cohesion can query AI tools for suggestions on transitions between sentences, while a student in need of syntactic variety can look for assistance on changing sentence structures. To address varying levels of student ability, those who require less scaffolding can be encouraged to use open-ended prompts, which promote greater autonomy and critical thinking. In contrast, students who need more support benefit from structured templates and step-by-step guidance. Such scaffolded prompts provide essential guidance at early stages and can be gradually withdrawn as students gain confidence and proficiency, ensuring that support is tailored to individual learning needs.

In addition to differentiated scaffolding, the development of AI literacy should be explicitly integrated into instruction. Teachers should guide students to not only receive AI-generated feedback, but also to critically evaluate its quality, relevance, and appropriateness for their own writing contexts. This can be achieved through classroom discussion, peer review, or reflective journals in which students are asked to compare AI suggestions with assignment rubrics, teacher feedback, or their own writing goals. Explicit training in AI literacy empowers learners to discern when to accept, modify, or reject AI suggestions, fostering deeper engagement and avoiding uncritical acceptance. Finally, it is critical to prepare students to interpret and integrate feedback from AI with an awareness of potential ethical concerns. Over-reliance on AI-generated feedback may lead to superficial learning or decreased confidence in students' own judgement. There is also a risk that students may come to rely too heavily on AI for editing, at the expense of developing independent writing and revision strategies. To address these concerns, teachers should emphasize the complementary role of AI in supporting—rather than replacing—human judgement, encourage active reflection, and maintain a balance between technology-mediated and teacher-guided feedback.

5. Conclusion

This paper examines Chinese university students' attitudes toward AI-assisted Written Corrective Feedback (AI-WCF) and their perceived achievement in academic writing. Findings, based on a retrospective pretest–posttest design and on an experience-based attitude survey, demonstrate that in general, students had positive attitudes towards AI-generated feedback. They reported significant improvements in writing knowledge and skills, especially in thematic unity, paragraph structure, cohesion, and lexical accuracy. The effect sizes of the gains over most variables were large, demonstrating both statistical significance and practical significance.

These results provide evidence that AI-WCF can be effective when situated in well-structured instructional contexts in support of writing development and learner autonomy. It also indicates the pedagogic value of integrating automated feedback with reflective strategies of learning, and in particular when combined with tailored scaffolding according to students' requirement.

Despite these promising results, several limitations should be noted. The sample was drawn from a single university and relied on self-reported data, which may affect the generalizability and objectivity of the findings. Although the retrospective pretest–posttest (RPP) design addresses response-shift bias, it may still be influenced by participants' expectations and the novelty of AI technology. The absence of a comparative control group and limited qualitative data also constrain the scope of interpretation.

Future research should employ mixed-methods designs, incorporate appropriate control groups, and include longitudinal follow-ups to examine the sustained impact and relative effectiveness of AI-WCF. Triangulating self-reports with objective measures and richer qualitative data will further clarify how students engage with and benefit from AI-generated feedback in diverse educational contexts.

Acknowledgments

Not applicable.

Authors contributions

The author was solely responsible for the study design, data collection, data analysis, and manuscript preparation. The author read and approved the final manuscript.

Funding

Not applicable.

Competing interests

The author declares no competing interests.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research, 3*(2), 818-824. https://doi.org/10.25082/AMLER.2023.02.009

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.

Bray, J. H., Maxwell, S. E., & Howard, G. S. (1984). Methods of analysis with response-shift bias. *Educational Measurement and Psychological Measurement, 44*, 781-804. https://doi.org/10.1177/0013164484444002

Chen, Z., Wei, W., Zhu, X., & Yao, Y. (2025). Unpacking the rejection of L2 students toward ChatGPT-generated feedback: An explanatory research. *ECNU Review of Education, 0*(0). https://doi.org/10.1177/20965311241305140

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97-107. https://doi.org/10.1093/elt/ccn023

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*(2), 181-201. https://doi.org/10.1017/S0272263109990490

Gardner, R. C., Lalonde, R. N., & Moorcroft, R. (1985). The role of attitudes and motivation in second language learning: Correlational and experimental considerations. *Language Learning, 35*(2), 207-227. https://doi.org/10.1111/j.1467-1770.1985.tb01025.x

Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *arXiv preprint* arXiv:2403.20216. https://doi.org/10.48550/arXiv.2403.20216

Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System, 126*, 103502. https://doi.org/10.1016/j.system.2024.103502

Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement, 3*, 1-23. https://doi.org/10.1177/014662167900300101

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

Jiang, Y. (2025). Interaction and dialogue: Integration and application of artificial intelligence in blended mode writing feedback. *The Internet and Higher Education, 64*, 100975. https://doi.org/10.1016/j.iheduc.2024.100975

Li, X. (2024). Teaching listening for interpreting through mind mapping: Students' attitudes and its effectiveness. *Revista Española de Lingüística Aplicada* [Spanish Journal of Applied Linguistics], *37*(1), 172-199. https://doi.org/10.1075/resla.21037.li

Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System, 127*, 103529. https://doi.org/10.1016/j.system.2024.103529

Little, T. D., et al. (2019). The retrospective pretest–posttest design redux: On its validity as an alternative to traditional pretest–posttest measurement. *International Journal of Behavioral Development, 43*(2), 173-180. https://doi.org/10.1177/0165025419877973

Osawa, K. (2023). Integrating automated written corrective feedback into e-portfolios for second language writing: Notion and Notion AI. *RELC Journal, 55*(3), 881-887. https://doi.org/10.1177/00336882231198913

Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning, 31*(7), 653-674. https://doi.org/10.1080/09588221.2018.1428994

Ren, Y., Liu, Y., & Xie, Y. (2024). 英语写作教学中 ChatGPT 与教师反馈的对比研究 [A comparative study of ChatGPT and teacher feedback in English writing instruction]. *外语教学理论与实践*, (4), 38-40. (in Chinese)

Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine, 48*(11), 1507-1515. https://doi.org/10.1016/S0277-9536(99)00045-3

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. https://doi.org/10.3390/educsci15020232

Wang, D. (2024). Teacher-versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *International Review of Research in Open and Distributed Learning, 25*(3), 37-56. https://doi.org/10.19173/irrodl.v25i3.7646

Wei, S., & Li, L. (2023). 人工智能辅助二语写作反馈研究——以 ChatGPT 为例 [A study of AI-assisted feedback in second language writing: A case of ChatGPT]. *中国外语, 20*(3), 33-40. (in Chinese)

Wiboolyasarin, W., Wiboolyasarin, K., Suwanwihok, K., Jinowat, N., & Muenjanchoey, R. (2024). Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments. *Computers and Education: Artificial Intelligence, 6*, 100228. https://doi.org/10.1016/j.caeai.2024.100228

Ziqi, C., Xinhua, Z., Qi, L., & Wei, W. (2024). L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. *Computer Assisted Language Learning*, 1-21. https://doi.org/10.1080/09588221.2024.2422478