Journal of Posthumanism
2025
Volume: 5, No: 8, pp. 293–300
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i8.3159

From Dependence to Development: The Role of AI in Enhancing Writing Proficiency Among Chinese English Majors

Hailong Zhang[1], Chunming Chen[2], Yanyan Liu[3], Liuqi Peng[4]

Abstract

The integration of artificial intelligence (AI) tools into higher education writing instruction is reshaping traditional pedagogical practices, yet empirical evidence on their effectiveness remains limited among English majors. This study aimed to examine the impact of human-AI collaborative instruction on the writing proficiency of English majors in a Chinese university. Using a quasi-experimental design, 80 third-year English majors were assigned to either an experimental group (AI-assisted instruction with teacher guidance) or a control group (traditional instruction). Writing proficiency was measured through pre- and post-intervention IELTS writing tasks, scored by three trained raters using analytic band descriptors. Findings revealed that the experimental group achieved significantly higher gains in overall writing performance compared to the control group ($p < .001$), with particularly large improvements in lexical resource and grammatical accuracy. ANCOVA results confirmed the instructional approach as a significant predictor of post-test performance, independent of baseline scores. The findings highlight the pedagogical potential of human-AI collaboration in improving writing outcomes and offer empirical support for its integration into curriculum design of English major.

Keywords: *AI-Assisted Writing, Human-AI Collaboration, English Majors, Writing Proficiency.*

Introduction

Universities worldwide are endeavouring to incorporate AI into their curricula, and the promise of AI-assisted writing education has attracted interest from researchers (Cardon et al., 2023; Chen & Gong, 2022). In the realm of Chinese higher education, English majors have distinct obstacles in enhancing their writing proficiency. Despite extensive formal language education, they frequently have difficulties with lexical variety, grammatical precision, and rhetorical structure (Pryma et al., 2025). Conventional writing education, marked by teacher-centric feedback and minimal individualised assistance, has demonstrated inadequacy in addressing the varied requirements of a substantial student demographic. Conventional methods frequently include protracted feedback loops, broad error rectification, and restricted chances for iterative refinement, resulting in sluggish learning advancement and enduring linguistic shortcomings (Nguyen et al., 2020).

The use of AI technology into writing teaching presents a viable solution to these constraints. AI-driven writing tools provide distinct benefits: instant feedback, tailored error identification, vocabulary enhancement recommendations, and round-the-clock access to student practice

[1] Henan Jinnan International Trade Co., Ltd. Zhengzhou, China, Email: zhanghailong.student@outlook.com
[2] School of Electronic Information Engineering, Guangzhou City Polytechnic, Guangzhou, China, Email: sharedemali@outlook.com
[3] Ningxia Shizuishan Industry and Trade Vocational TechAcademy, Shizuishan, China, Email: 664612251@qq.com
[4] Department of English Language Education, The Education University of Hong Kong, Hong Kong SAR, China, Email: Iqpeng@eduhk.hk, (Corresponding Author)

(Zhao, 2024). Human-computer cooperation in education signifies a paradigm change, transitioning from perceiving technology as a substitute for or as a supplementary tool for human instruction. This technique highlights the synergistic potential of integrating human pedagogical experience with AI's analytical skills to enhance learning settings (Alshahrani & Qureshi, 2022). This partnership in writing teaching may entail AI delivering instantaneous language feedback, allowing educators to concentrate on advanced issues such as argumentation, critical thinking, and rhetorical efficacy.

Nonetheless, despite increasing interest for AI applications in education, empirical data about its efficacy in enhancing writing abilities is few, especially among English majors in higher education contexts. Although several studies have investigated the use of AI in the context of general English as a foreign language writing (Hwang, 2023), there is a paucity of research rigorously analysing the effects of human-computer collaborative education on the particular writing difficulties encountered by English majors. This study investigates the efficacy of human-computer collaborative writing training for English majors at Chinese colleges, addressing existing deficiencies. It investigates whether, when integrated with conventional teacher education, it may substantially enhance writing proficiency across several aspects of academic writing competence. This study utilises a quasi-experimental methodology and recognised assessment frameworks to furnish empirical information about the pedagogical efficacy of human-computer cooperation in higher education writing training. The ramifications of this study transcend its local institutional framework.

Literature Review

Recent empirical research has started to demonstrate the efficacy of AI-assisted writing teaching in several educational settings. Studies investigating ChatGPT's influence on English as second language (ESL) students' academic writing abilities have yielded encouraging outcomes. The results demonstrate a substantial favourable influence of ChatGPT on students' academic writing abilities, and students' evaluations of this impact were predominantly favourable (Lingard, 2023). Studies also investigate automatic writing evaluation among Chinese ESL learners has shown favourable results. In the expanding domain of second language (L2) education, where adept writing is crucial for successful language learning, research indicates that AWE technologies can offer substantial advantages to student writers (Zhang & Tsung, 2021; Li, 2020).

A thorough investigation into ChatGPT's effectiveness in AI-supported language acquisition for EFL students shown significant enhancements. Quantitative study indicates substantial enhancements in writing proficiency and motivation among students who underwent AI-assisted instruction relative to the control group (Song & Song, 2023). The experimental group exhibits improved competence in several facets of writing, such as organisation and coherence.

Research has identified specific areas where AI tools demonstrate particular effectiveness in supporting student writing. AI enhances academic writing in six areas: idea generation, content structuring, literature synthesis, data management, editing, and ethical compliance (Amirjalili et al., 2024; Nazari et al., 2021). This comprehensive support suggests that AI can assist students throughout the entire writing process. Positive effects were reported on students' psychological factors related to writing. After being exposed to computer-based feedback, Sherafati and Mahmoudi Largani's (2022) study found that Iranian ESL learners' writing self-efficacy scale scores improved significantly, indicating that AI feedback can enhance students' confidence in their writing abilities.

Research on academic writing English's (AWE) impact on peer and self-editing has yielded encouraging results (AI-Inbari et al., 2023). The AWE software has positively affected both the peer and self-editing students (Chen & Cui, 2022). These studies reflect students' positive evaluation of both the software and the automated feedback provided.

Methodology

Research Design

This study adopts a quasi-experimental design to investigate the impact of human-AI collaborative writing instruction on the writing proficiency of English majors in a Chinese university context. The study follows a pre-test/post-test non-equivalent groups design, involving two intact classes from the same English department. One class was assigned as the experimental group, receiving writing instruction integrated with AI support tool, DeepSeek, alongside teachers' guide. The other class is control group, receiving conventional writing instruction without the AI technologies. Pre- and post-intervention writing assessments were administered to both groups to evaluate changes in writing performance across four dimensions: content development, organization, vocabulary usage, and grammatical accuracy.

Participants

The participants of this study consisted of 80 undergraduate students majoring in English at a public university in eastern China. All participants were third-year students enrolled in the required Academic English Writing course, which is offered as part of the university's core curriculum for English majors. These students had completed at least four semesters of English language and writing instruction prior to the study, ensuring a relatively homogenous level of foundational proficiency.

Participants were recruited from two classes, with one class (n = 40) designated as the experimental group and the other (n = 40) as the control group. The assignment was based on pre-existing class arrangements, and no randomization was conducted to preserve the authenticity of the instructional environment. However, pre-test results were used to confirm baseline equivalence in writing proficiency between the two groups, thereby minimizing potential selection bias. The demographic profile of the participants was balanced across groups, with approximately 72% female and 28% male students. Participants' ages ranged from 20 to 22 years. All students voluntarily participated in the study and provided informed consent prior to data collection.

Instruments

This study employed multiple data sources, including writing assessments, student writing samples, and semi-structured interviews with instructors. All instruments were selected or designed to ensure alignment with the research objectives and to enhance the reliability and validity of the findings.

A pre-test and post-test writing task were administered to both the experimental and control groups at the beginning and end of the 8-week intervention. The prompts were parallel in topic and difficulty and were designed to elicit academic argumentative writing. Each writing task was limited to 250–300 words and completed within a 60-minute in-class session without access to external assistance or AI tools. The writing samples were evaluated using an analytic scoring rubric adapted from the IELTS Writing Band Descriptors, including four dimensions: Task Response (TR), Coherence and Cohesion (CC), Lexical Resource (LR), and Grammatical Range

and Accuracy (GRA).

Each dimension was rated on a 0–9 scale, with three independent raters (experienced writing instructors) scoring all scripts. Inter-rater reliability was established through a pilot scoring session, yielding a Cohen's kappa coefficient of 0.82, indicating strong agreement.

Data Analysis

To examine the impact of human-AI collaborative writing instruction on students' writing performance, descriptive statistics and inferential tests were conducted using SPSS 27.0. First, descriptive statistics (means and standard deviations) were calculated for both the pre-test and post-test writing scores across the experimental and control groups. A Shapiro-Wilk test was used to assess the normality of the score distributions. Given that the assumptions of normality and homogeneity of variance were met, a paired-samples t-test was used to determine whether there was a significant improvement in writing scores within each group.

To assess between-group differences while controlling for baseline performance, an Analysis of Covariance (ANCOVA) was performed, using the pre-test score as the covariate and the post-test score as the dependent variable. This allowed for the evaluation of instructional effects attributable to the AI-human collaboration intervention, independent of students' initial writing ability. The significance level was set at $p < 0.05$, and effect sizes (Cohen's d and partial eta squared) were reported to assess the magnitude of the observed differences.

Findings

Descriptive Statistics

To provide a more nuanced picture of students' writing performance, descriptive statistics were calculated across the four IELTS writing dimensions. Table 1 displays the mean scores and standard deviations for each dimension in both pre- and post-tests for the experimental and control groups. As seen in Table 1, the experimental group demonstrated noticeable improvements across all four writing dimensions, particularly in LR (+0.93) and GRA (+0.64). These gains suggest that the use of DeepSeek was especially effective in helping students identify and correct lower-level language issues. In contrast, the control group exhibited marginal improvements in all four dimensions, with average gains ranging from +0.15 to +0.29. The relatively limited improvement may reflect the slower pace of learning through conventional instruction.

Group	Test	TR (M ± SD)	CC (M ± SD)	LR (M ± SD)	GRA (M ± SD)	Overall Mean
Experimental	Pre-test	5.74 (0.42)	5.65 (0.45)	5.58 (0.49)	5.75 (0.47)	5.68
	Post-test	6.38 (0.36)	6.41 (0.33)	6.51 (0.38)	6.39 (0.40)	6.42
Control	Pre-test	5.69 (0.43)	5.58 (0.44)	5.63 (0.46)	5.72 (0.41)	5.66
	Post-test	5.84 (0.38)	5.91 (0.39)	5.96 (0.41)	5.92 (0.36)	5.91

Table 1. Descriptive Statistics by IELTS Writing Dimensions

Within-Group Comparison

To determine whether the observed gains in writing proficiency were statistically significant within each group, paired-samples t-tests were conducted, comparing pre- and post-test scores for both the experimental and control groups. As shown in Table 2, the experimental group showed a statistically significant improvement in writing scores from pre-test to post-test, $t(39) = 11.53$, $p < .001$. The effect size (Cohen's $d = 1.82$) indicates a very large effect, suggesting that the AI-human collaborative writing instruction had a substantial impact on students' writing proficiency. The control group also showed a statistically significant but smaller improvement, $t(39) = 3.29$, $p = .002$, with a medium effect size (Cohen's $d = 0.52$). This result suggests that conventional instruction may lead to moderate gains over time, but not at the level observed in the experimental condition.

Group	t(df)	p-value	Effect Size (Cohen's *d*)
Experimental	11.53 (39)	.00***	1.82
Control	3.29 (39)	.02	0.52

Table 2. Paired-Samples t-Test Results

Between-Group Comparison

To further examine whether the observed post-test differences between the experimental and control groups were statistically significant after accounting for initial disparities in writing proficiency, ANCOVA was conducted. The post-test overall writing score (IELTS band, unrounded) was entered as the dependent variable, and the pre-test score served as the covariate. Group (experimental vs. control) was treated as the fixed factor. The ANCOVA revealed a statistically significant effect of instructional condition on post-test writing scores, $F(1, 77) = 21.82$, $p < .001$, after adjusting for pre-test scores. The partial η^2 of 0.221 indicates a large effect size, suggesting that the difference in post-intervention writing performance was not due to chance, but rather attributable to the instructional approach.

The adjusted mean post-test score for the experimental group was 6.41, compared to 5.92 for the control group, confirming the superior effectiveness of human-AI collaborative instruction in improving students' writing proficiency. This finding reinforces the descriptive and within-group t-test results, offering strong evidence that AI-assisted writing support, when integrated into a pedagogically guided framework, can produce measurable and educationally meaningful gains in writing performance.

Source	SS	df	MS	F	p-value	Partial η^2
Group (Instruction Type)	2.391	1	2.391	21.82	< .001	0.221
Pre-test (Covariate)	1.378	1	1.378	12.59	.001	0.141
Error	8.440	77	0.110			
Total	—	79	—			

Table 3. ANCOVA Summary for Overall Writing Score (Post-Test Adjusted by Pre-Test)

Subskill-Level Improvements

In addition to the overall gains in writing proficiency, further analysis was conducted to examine improvements across the four analytic dimensions of the IELTS writing rubric. These subskill-

level comparisons offer deeper insights into the specific areas influenced by the instructional interventions. As shown in Figure 1, students in the experimental group exhibited consistent gains across all four dimensions, with the most pronounced improvements in LR (+0.93) and GRA (+0.64). These findings suggest that AI-assisted tools were particularly effective in enhancing vocabulary usage and grammatical precision—two lower-level language skills that benefit from immediate, data-driven feedback. In contrast, the control group showed only marginal increases, with gains ranging from +0.15 (TR) to +0.29 (GRA). While this improvement may be attributed to ongoing writing practice and instructor feedback, the effect was significantly smaller than that observed in the experimental group.

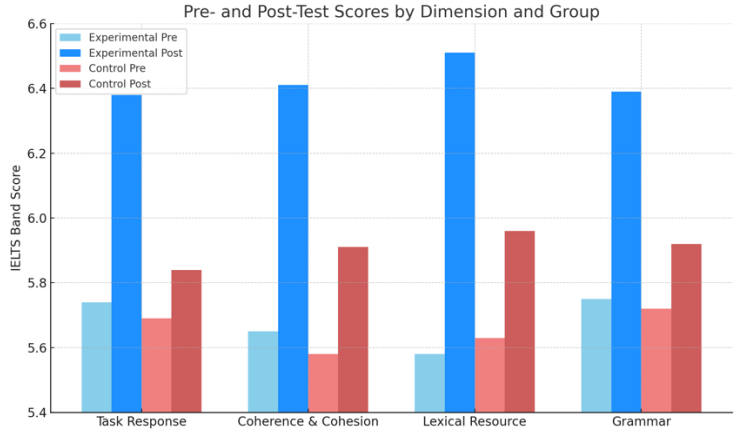

Figure 1. Pre- and Post-Test IELTS Scores by Dimension and Group

Discussion

This study aimed to investigate the impact of human-AI collaborative instruction on the writing proficiency of English majors in a Chinese university context. Drawing on rater-averaged IELTS writing scores, the findings revealed that students in the experimental group—who received AI-assisted instruction integrated with teacher guidance—achieved significantly greater improvements than those in the control group across both overall scores and specific subskills. The most notable gains were observed in lexical resource and grammatical range and accuracy, suggesting that the use of AI tools is particularly effective in enhancing lower-level language skills.

The significant increase in writing proficiency among the experimental group can be attributed to the synergistic effect of human-AI collaboration. AI tool provided immediate, personalized feedback on grammar and word choice, which may accelerate error correction and self-awareness during revision. When combined with structured teacher input—such as assignment scaffolding and goal-setting—students were likely able to engage in more deliberate and self-regulated writing practice, consistent with theories of self-regulated learning (Zimmerman, 2002).

The disproportionately large gains in Lexical Resource and Grammatical Range align with prior research suggesting that AI systems are particularly adept at flagging surface-level errors and suggesting lexical alternatives (Zhang & Zou, 2022). In contrast, the relatively modest improvement in Task Response indicates that idea development and argumentative quality—higher-order aspects of writing—still depend heavily on teacher facilitation, as AI tools are less

capable of evaluating rhetorical effectiveness in context.

These findings suggest that AI functions best as a complementary support system rather than a replacement for teacher feedback, particularly when deployed in targeted ways that align with learning goals.

Conclusion

The results offer several practical implications for English for Academic Purposes instructors and curriculum designers. Instructors could integrate AI tools into writing instruction can lead to measurable improvements, especially in areas that benefit from repeated, low-stakes feedback (e.g., grammar, word choice). Writing assignments should be designed to leverage AI for lower-order concerns, while preserving teacher authority in guiding high-level rhetorical and conceptual development.

This study contributes to the growing literature on AI-assisted learning by offering empirical support for a layered model of writing development, wherein AI supports lower-level linguistic accuracy, while teachers remain essential for guiding higher-order thinking and structure. The results validate key tenets of SRL theory, particularly in terms of goal-setting, self-monitoring, and feedback loops, and extend the TPACK framework by highlighting the dynamic interaction between technological tools and pedagogical scaffolding in writing education.

Limitations and Implications

Several limitations should be acknowledged. First, the study was conducted at a single institution with a relatively homogeneous sample of English majors, which may limit generalizability. Second, the intervention period was limited to eight weeks, and long-term retention of writing improvement was not assessed. Third, while rater reliability was established, future studies could benefit from textual corpus analysis or automated writing evaluation metrics to complement human scoring. Future research could explore comparative studies of different AI tools and their distinct instructional affordances.

References

Al-Inbari, F. A. Y., & Al-Wasy, B. Q. M. (2023). The impact of automated writing evaluation (AWE) on EFL learners' peer and self-editing. Education and Information Technologies, 28(6), 6645-6665.

Alshahrani, K., & Qureshi, R. J. (2024). Review the prospects and obstacles of AI-enhanced learning environments: the role of ChatGPT in education. International Journal of Modern Education and Computer Science, 16(4), 71-86.

Amirjalili, F., Neysani, M., & Nikbakht, A. (2024, March). Exploring the boundaries of authorship: A comparative analysis of AI-generated text and human academic writing in English literature. In Frontiers in Education (Vol. 9, p. 1347421). Frontiers Media SA.

Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. Business and Professional Communication Quarterly, 86(3), 257-295.

Chen, M., & Cui, Y. (2022). The effects of AWE and peer feedback on cohesion and coherence in continuation writing. Journal of Second Language Writing, 57, 100915.

Chen, C., & Gong, Y. (2025). The role of AI-assisted learning in academic writing: A mixed-methods study on Chinese as a second language students. Education Sciences, 15(2), 141.

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. Heliyon, 7(5).

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. Studies in Higher Education, 49(5), 847-864.

Hwang, W. Y., Nurtantyana, R., Purba, S. W. D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts. Journal of Educational Computing Research, 61(5), 1008-1035.

Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. Perspectives on medical education, 12(1), 261.

Li, C. (2020). A positive psychology perspective on Chinese EFL students' trait emotional intelligence, foreign language enjoyment and EFL learning achievement. Journal of Multilingual and Multicultural Development, 41(3), 246-263.

Pryma, V., Pelivan, O., Teletska, T., Tsobenko, O., & Zagrebelna, N. (2025). AI Writing Assistants and Student Competence: A Linguistic Aspect. Arab World English Journal, 16.

Sherafati, N., & Mahmoudi Largani, F. (2023). The potentiality of computer-based feedback in fostering EFL learners' writing performance, self-regulation ability, and self-efficacy beliefs. Journal of Computers in Education, 10(1), 27-55.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. Frontiers in psychology, 14, 1260843.

Zhang, L., & Tsung, L. (2021). Learning Chinese as a second language in China: Positive emotions and enjoyment. System, 96, 102410.

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: An analysis of language precision, content summarization, and creative writing facilitation. Education and Information Technologies, 1-32.