Generative AI Chatbot-Supported Process-Genre Writing: Chinese EFL Students' Writing and Self-Efficacy

Niu Chuhan[1]*, Omer Hassan Ali Mahfoodh[1], Nagaletchimee Annamalai[2,3]

[1] School of Languages, Literacies and Translation, Universiti Sains Malaysia, 11800 Pinang, Malaysia

[2] Department of Languages and Culture, College of Humanities and Sciences, Ajman University, Ajman, United Arab Emirates

[3] Humanities and Social Sciences Research Center (HSRC), Ajman University, Ajman, United Arab Emirates

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.101100273

Received: 01 November 2026; Accepted: 06 November 2026; Published: 25 November 2026

ABSTRACT:

Generative artificial intelligence (AI) has become increasingly prominent in language education, yet its pedagogical integration into established writing approaches remains underexplored. The present study integrates a generative AI chatbot into the process-genre writing approach, conceptualising AI as a pedagogical scaffold rather than merely a writing tool. Specifically, this study investigated the effects of a generative AI chatbot-supported process-genre writing approach on Chinese university students' argumentative writing performance and writing self-efficacy in a blended learning environment. An explanatory sequential mixed-methods design was employed with 69 first-year university students assigned to an experimental group (n = 34) and a control group (n = 35). The experimental group received chatbot-supported process-genre writing instruction, whereas the control group received conventional writing instruction. Data were collected through pre- and post-writing tests, a writing self-efficacy questionnaire, and semi-structured interviews. Quantitative data were analysed using one-way ANCOVA, paired-samples t-tests, and independent sample t-tests, while qualitative data were analysed using thematic analysis. The experimental group demonstrated significantly greater improvements in argumentative writing performance and all dimensions of writing self-efficacy than the control group. Analysis of the interviews revealed that the chatbot supported planning, drafting, revision, and reflection while providing immediate feedback and affective support, although concerns regarding over-reliance and loss of individual voice were also identified. The findings demonstrate the value of integrating generative AI as a pedagogical scaffold within process-genre writing instruction.

Keywords: Generative AI chatbot, process-genre writing, argumentative writing, writing self-efficacy, blended learning

INTRODUCTION

Argumentative writing is a fundamental academic literacy skill that enables learners to construct evidence-based arguments, evaluate competing perspectives, and communicate effectively in academic and professional contexts (Kuhn & Moore, 2015; Zhu, 2001). However, it remains one of the most challenging genres for English as a Foreign Language (EFL) learners because of its cognitive, linguistic, and rhetorical demands (Kerman et al., 2024; Yang, 2020). Producing effective argumentative texts requires learners to develop logical arguments, organise ideas coherently, engage in dialogic reasoning, use appropriate language, and refine their writing through feedback and interaction (Chin & Osborne, 2010; Guo et al., 2024; Su et al., 2023; Su et al., 2021). These demands often weaken learners' confidence in writing, making writing self-efficacy a key factor in writing development. Writing self-efficacy, defined as learners' beliefs in their capability to perform writing tasks successfully (Bandura, 1997b), has consistently been associated with greater motivation, lower anxiety, stronger

persistence, and better writing performance (e.g., Grigg et al., 2018; Kong & Teng, 2023; Schunk & DiBenedetto, 2020; Woodrow, 2011). Because writing is a recursive process involving planning, drafting, revising, and monitoring (e.g., Hayes, 2012; Pajares, 2003; Teng et al., 2018; Wang et al., 2024), effective instruction should enhance both writing performance and writing self-efficacy.

Recent advances in generative artificial intelligence (AI) have created new opportunities for L2 writing instruction. Earlier research focused mainly on automated writing evaluation systems, such as Grammarly and Pigai.org, which provide corrective feedback after texts are produced (Li & Yan, 2020; Wei et al., 2023). More recently, generative AI has evolved into an interactive writing assistant capable of supporting idea generation, revision, and dialogue throughout the writing process (e.g., Jin, 2025; Lai et al., 2026; Oppenheimer et al., 2025; Pellas, 2023; Ruwe & Mayweg-Paus, 2023). Although previous studies generally report improvements in writing performance, engagement, and motivation, findings regarding writing self-efficacy remain inconsistent, with some studies reporting positive effects and others identifying declines in learners' confidence (e.g., R. Zhang et al., 2023; Zhang et al., 2025). Moreover, most studies have examined AI as a writing assistant or feedback provider rather than investigating how it can be systematically embedded within an established pedagogical framework. Consequently, limited evidence exists on how generative AI can function as a pedagogical scaffold throughout the writing process, influence multidimensional writing self-efficacy, and support learning in blended environments that combine teacher guidance, peer interaction, and out-of-class AI use.

To address these gaps, the present study integrates a generative AI chatbot into the process-genre writing approach and conceptualises AI as a pedagogical scaffold rather than merely a technological tool. Using an explanatory sequential mixed-methods design, the study examines its effects on Chinese university EFL students' argumentative writing performance and five dimensions of writing self-efficacy, while qualitative interviews explore students' learning experiences. By demonstrating how generative AI can be systematically integrated into an established instructional framework, this study extends current understanding of AI-supported writing pedagogy in blended learning environments. Accordingly, the study addresses the following research questions:

1. What is the effect of a generative AI chatbot-supported process-genre writing approach in a blended learning environment on Chinese EFL students' argumentative writing performance?

2. To what extent do Chinese EFL students who receive a generative AI chatbot-supported process-genre writing approach demonstrate greater improvement in writing self-efficacy than those who receive conventional instruction in a blended learning environment?

3. How do Chinese EFL students perceive the use of generative AI chatbots throughout the process-genre writing process in a blended learning environment?

LITERATURE REVIEW

AI technology in L2 writing

AI, particularly chatbot technology, has become an increasingly important innovation in language education because of its potential to provide immediate, interactive, and individualized learning support. Previous studies have demonstrated that chatbots can facilitate second language acquisition by supporting language practice, formative assessment, learner engagement, motivation, and a wide range of language skills (e.g., Ebadi & Amini, 2024; Jeon, 2023; Kohnke, 2023). Within English as a Foreign Language (EFL) context, chatbots have been successfully employed to improve vocabulary acquisition, speaking, grammar, and reading comprehension (Chen et al., 2020; Han, 2020). Despite these promising developments, chatbot integration into L2 writing instruction has received considerably less attention than other language skills (Huang et al., 2022), and the pedagogical role of AI in supporting writing development remains insufficiently understood (Yan, 2023).

Recent advances in large language models have transformed AI from an automated evaluation tool into an interactive learning partner capable of supporting learners throughout the writing process. Earlier AI applications primarily functioned as writing evaluators that provided corrective feedback on grammar, vocabulary, and language accuracy after learners had completed their texts. In contrast, generative AI can engage learners in

iterative dialogue, generate alternative ideas, explain concepts, provide individualized feedback, and participate in collaborative writing activities (Hartwell & Aull, 2022). These capabilities are particularly valuable in argumentative writing, where learners require support not only in linguistic accuracy but also in generating ideas, constructing logical arguments, organizing evidence, and revising their texts. Consequently, an increasing number of studies have investigated how generative AI supports different stages of writing, including pre-writing, drafting, revising, and post-writing activities (e.g., Dong, 2024; Guo et al., 2023; Mahdi et al., 2025; S. Zhang et al., 2023). Collectively, these studies suggest that AI-assisted writing can improve feedback quality, revision practices, language accuracy, and overall writing performance.

Although these findings demonstrate the educational potential of generative AI, important limitations remain. Most previous studies have examined relatively short-term interventions, isolated writing stages, or individual writing tasks, with limited attention given to how AI can be systematically integrated within an established instructional approach. In many cases, generative AI has been treated primarily as a writing assistant, feedback provider, or virtual peer rather than as a pedagogically designed instructional scaffold embedded throughout the writing process. Consequently, there remains limited empirical evidence regarding how chatbot-supported instruction can enhance content development, organization, language use, and writing self-efficacy simultaneously within a structured process-genre approach. Furthermore, relatively little research has examined AI-supported writing instruction in blended learning environments where teacher guidance, peer collaboration, classroom instruction, and out-of-class AI interaction operate together. Addressing these gaps is important for understanding not only whether generative AI improves writing, but also how it can be pedagogically integrated to support sustainable writing development in higher education.

Review of related studies

A synthesis of the reviewed studies indicates that generative AI has considerable potential to support L2 writing by enhancing writing performance, engagement, and learners' confidence. Intervention studies have generally reported positive effects on writing quality, writing self-efficacy, or both following AI-supported instruction, although the magnitude and consistency of these improvements have varied across educational contexts, learner populations, and instructional designs (e.g., Can, 2026; Gao et al., 2025; Lai et al., 2026; Li & Tekwa, 2025; Zhang et al., 2025). In contrast, some recent studies have adopted cross-sectional or correlational designs that examine learners' perceptions of AI, AI acceptance, AI-assisted writing self-efficacy, or AI-use self-efficacy rather than investigating changes resulting from instructional interventions (e.g., Almusharraf et al., 2025; Jin, 2025; Shen & Tao, 2025; Xue & Chen, 2026; Yao & Fan, 2026; Zhou & Wang, 2026). All these findings suggest that generative AI has considerable potential to facilitate both cognitive and affective aspects of L2 writing, while also highlighting the need for more robust intervention-based research.

Despite these advances, important gaps remain in the literature. Existing intervention studies have primarily examined generative AI as a writing assistant, feedback provider, or task-support tool, with limited attention to its systematic integration within established writing pedagogies. Although some studies incorporated teacher guidance alongside AI, few conceptualised AI as a pedagogical scaffold embedded throughout the writing process or examined its combined influence on argumentative writing performance and writing self-efficacy within a blended learning environment. Furthermore, relatively few studies have adopted explanatory sequential mixed-methods designs to explain not only whether AI-supported instruction is associated with improved learning outcomes but also how students experience AI-supported writing throughout the instructional process. Addressing these gaps can provide a more comprehensive understanding of the pedagogical role of generative AI in EFL writing instruction.

THEORETICAL FRAMEWORK

The present study is grounded in the cognitive process theory of writing developed by Flower and Hayes (1981) and the social cognitive theory proposed by Bandura (1997a). Together, these theories explain how chatbot-supported process-genre writing instruction may influence both argumentative writing performance and writing self-efficacy. Flower and Hayes (1981) conceptualized writing as a recursive process involving planning, translating, reviewing, and monitoring. Building on this perspective, the process-genre approach guides learners

through genre analysis, planning, drafting, revising, editing, proofreading, and reflection. Within this framework, the generative AI chatbot functions as a pedagogical scaffold rather than a substitute writer by supporting idea generation, argument development, organization, language refinement, and reflection, while teacher guidance and peer interaction encourage learners to critically evaluate AI-generated suggestions and maintain ownership of their writing. Complementing this perspective, social cognitive theory explains how these learning experiences strengthen writing self-efficacy through mastery experiences, vicarious learning, social persuasion, and affective states Bandura (2012). Accordingly, these two theories provide the conceptual foundation for examining how chatbot-supported process-genre instruction influences students' argumentative writing performance (RQ1), writing self-efficacy (RQ2), and learning experiences (RQ3). Together, they position generative AI not merely as a technological writing tool but as a pedagogically embedded instructional scaffold that supports both the cognitive and affective dimensions of argumentative writing development.

METHOD

Research design and participants

This study employed an explanatory sequential mixed-methods design with a quasi-experimental approach to examine the effects of a generative AI chatbot-supported process-genre writing approach on Chinese EFL students' argumentative writing performance and writing self-efficacy in a blended learning environment. The quantitative phase employed pre-test and post-test writing assessments together with pre- and post-intervention writing self-efficacy questionnaires to evaluate the effectiveness of the instructional intervention. The subsequent qualitative phase consisted of semi-structured interviews that explored students' learning experiences and explained the quantitative findings by providing deeper insights into how chatbot-supported instruction facilitated or constrained their writing development.

The study involved 69 first-year non-English major students enrolled in a General English course at a Chinese university. As students had already been assigned to classes before the study, random assignment of individuals was not feasible. Therefore, the two intact classes were randomly assigned to either the experimental group (n = 34) or the control group (n = 35), making a quasi-experimental design appropriate. The experimental class received chatbot-supported process-genre writing instruction, whereas the control class received conventional instruction involving model-text analysis, imitation, guided writing practice, and teacher feedback. Both classes completed the same four writing topics and received the same amount of scheduled instructional time. The participants had a mean age of 19.5 years (30 males and 39 females), and both groups demonstrated comparable English proficiency before the intervention. Each class was taught by a different instructor. To ensure instructional consistency, both instructors received training on the instructional procedures for their respective conditions to support the implementation of the planned teaching activities. Both classes were taught by experienced instructors who followed standardized teaching procedures. Participation was voluntary, written informed consent was obtained from all participants, and anonymity was maintained throughout the study.

Data collection

Data collection followed an explanatory sequential mixed-methods design. In Week 1, all 69 participants completed the argumentative writing pre-test and the pre-intervention writing self-efficacy questionnaire under supervised classroom conditions without assistance from teachers, peers, or AI tools. In Week 2 to Week 9, the experimental group received chatbot-supported process-genre writing instruction, whereas the control group received conventional writing instruction while completing four argumentative writing tasks. In Week 10, both groups completed the post-test and post-intervention questionnaire. The pre-test and post-test followed the same time limit and word-count requirement: students were required to write an argumentative essay of more than 200 words within 45 minutes. The two tests used different topics selected to be similar in difficulty. Students were not permitted to use AI tools or receive assistance from teachers or peers during either assessment. All 69 participants, comprising 34 students in the experimental group and 35 in the control group, completed both writing tests and both administrations of the self-efficacy questionnaire. No missing responses or excluded cases were identified, and all participants were included in the quantitative analyses. Based on the quantitative findings, nine students from the experimental group were purposively selected for individual semi-structured

interviews to provide deeper insights into their learning experiences and explain the quantitative results. As explained below, three instruments were used to collect data.

Argumentative writing tests

Argumentative writing tests were used to compare the writing performance of the experimental and control groups. Students were required to write an argumentative essay of more than 200 words within 45 minutes. The essays were assessed using the argumentative writing rubric developed by Li et al. (2025). This rubric was selected because it was developed and validated specifically for assessing EFL argumentative writing, making it well directly related with the genre and learner population of the present study. Drawing on the assessment criteria proposed by Pagano et al. (2008), Awada et al. (2020), and Chang et al. (2024), the rubric assesses three dimensions including content, structure, and language. Content focused on argument relevance, the sufficiency and persuasiveness of supporting evidence, and the logical connections between evidence and claims. Structure assessed the completeness and logical organization of the essay, paragraph coherence, and transitions. Language assessed accuracy in vocabulary, grammar, and punctuation, alongside clarity, conciseness, and readability. Each dimension was rated on a 10-point scale using three performance bands (1–4, 5–7, and 8–10), reflecting progressively stronger performance. Content, structure, and language were weighted at 40%, 40%, and 20%, respectively, yielding a weighted total score on a 10-point scale.

Two raters independently assessed the essays. Both held doctoral degrees and had backgrounds in English education, with each having 10 years of teaching experience and eight years of experience assessing CET-4 and CET-6 writing. Before formal scoring, they reviewed the rubric, practiced scoring sample essays, and discussed the assessment criteria to develop a shared understanding of the scoring standards. The raters were unaware of students' group membership and whether the essays were from the pre-test or post-test. When the two raters' weighted overall scores differed by more than two points on the 10-point scale, a third rater was consulted to adjudicate the discrepancy.

Self-efficacy questionnaire

The writing self-efficacy questionnaire was adapted from the 27-item Questionnaire of English Writing Self-Efficacy (QEWSE) examined by Sun et al. (2022). The adapted version contained 25 items addressing students' confidence in English argumentative writing. Original Items 14 and 17, concerning email and diary writing, were removed. Other items were retained or modified to reflect argumentative writing activities, with changes to both item content and order. The adapted questionnaire included items addressing the development of introductions, the consideration of counterarguments, the writing of conclusions, and the use of an academic tone. The original seven-point response scale was retained, ranging from 1 ("I cannot do it at all") to 7 ("I can do it well"). These changes were improved and validated through two experts' evaluation. These experts with PhD Degree in TESOL examined the questionnaire items for readability, clarity, and comprehensiveness and suggest modifications if any. Their feedback resulted in improving the wording and clarity of items. The original seven-point Likert scale was retained. The adapted English questionnaire was translated into Chinese by a bilingual language expert. A second bilingual expert then reviewed the Chinese version against the English version to check the accuracy of the translation and consistency of meaning. The adapted questionnaire was piloted with three students of comparable English proficiency who did not participate in the main study. After completing the questionnaire, they were asked whether any items were difficult to understand, ambiguous, or difficult to answer, and were invited to identify unclear words or sentences. None of the students identified unclear wording, and no further wording revisions were made following the pilot.

Based on the QEWSE, the adapted questionnaire comprised 25 items relevant to ideation, organization, grammar and spelling, use of English writing, and self-efficacy for self-regulation. The original instrument developed by Sun and Wang (2020) was validated with Chinese university students and demonstrated good internal consistency (Cronbach's $\alpha = .94$ for the overall scale, .71 for ideation, .82 for organization, .78 for grammar and spelling, .87 for use of English writing, and .78 for self-regulation). In the present study, the adapted questionnaire also demonstrated good internal consistency, with Cronbach's alpha coefficients of .95 for the overall scale, .76 for ideation, .86 for organization, .80 for grammar and spelling, .77 for use of English writing,

and .78 for self-regulation. These findings indicate that the questionnaire provided reliable measurements of the participants' writing self-efficacy.

Semi-structured interviews

Semi-structured interviews were conducted to explore students' experiences of chatbot-supported process-genre writing instruction in the blended learning environment and to provide qualitative explanations for the quantitative findings. The interview protocol was reviewed by the course instructor and two students for clarity and relevance before being pilot-tested with two students who were not included in the main study. Individual interviews were chosen instead of focus groups to encourage open discussion and minimise the possibility that lower-proficiency students would feel reluctant to express their views in the presence of higher-proficiency peers. Participants were selected using stratified purposive sampling. Based on the post-test argumentative writing scores, the experimental group was divided into three proficiency levels (low, moderate, and high), and three students were purposively selected from each level, resulting in a total of nine participants. This sampling strategy ensured representation across different writing proficiency levels. Interviews were conducted concurrently with data analysis, and thematic saturation was reached after the seventh interview. Two additional interviews were conducted to confirm that no substantially new themes emerged.

The interviews were conducted individually by two trained PhD holders and lasted approximately 45 minutes each. All interviews were conducted in Chinese to enable participants to express their experiences naturally and in greater depth. With participants' consent, the interviews were audio-recorded and transcribed verbatim in Chinese. The Chinese transcripts served as the primary data for analysis, while selected quotations were translated into English for reporting. To ensure translation accuracy, a second bilingual expert independently compared the English translations with the original Chinese transcripts, and any discrepancies were resolved through discussion before the quotations were included in the manuscript.

Intervention Program

The eight-week intervention comprised four argumentative writing tasks. Both groups completed the same writing topics and received the same amount of scheduled instructional time. The experimental group used ERNIE Bot within a chatbot-supported process-genre writing programme that combined classroom instruction with out-of-class practice. ERNIE Bot is a generative AI chatbot developed by Baidu that responds to natural-language prompts and supports text generation and content creation (Baidu, Inc., 2023). In the present study, it was used to support students' argumentative writing through idea generation, the organisation of arguments and counterarguments, and feedback on content, organisation, and language. The instructional prompts were adapted from OpenAI's A Student's Guide to Writing with ChatGPT (OpenAI, 2024) for use with ERNIE Bot in the argumentative writing tasks. The adapted prompts were subsequently reviewed by two experts in artificial intelligence. During each task, students analyzed model texts, identified argumentative genre features, and participated in teacher-guided activities and peer discussions. Chatbot support was incorporated into idea generation, the organization of arguments and counterarguments, revision of content and organization, language refinement, and reflection on writing. The teacher demonstrated prompting strategies and guided students in critically evaluating chatbot responses, including identifying inaccurate or inappropriate suggestions. The teacher also monitored chatbot interactions, and students were encouraged to continue using the chatbot outside class to revise their drafts. These activities positioned the chatbot as a source of scaffolding throughout the writing process. The control group completed the same four tasks without chatbot support. Instruction consisted of teacher explanations, model-text analysis and imitation, guided writing practice, and conventional teacher feedback on content, organization, and language. The comparison therefore examined chatbot-supported process-genre instruction in relation to conventional writing instruction.

Data analysis

The quantitative data were analysed using IBM SPSS Statistics 27. The argumentative essays were independently scored by two trained raters using the established writing rubric, and the mean of the two ratings was used as the final score. Inter-rater reliability was evaluated using Cohen's kappa and the intraclass correlation coefficient (ICC), indicating acceptable agreement between the raters. Before analysis, the data were screened for missing

values and tested for normality (Pallant, 2020). For RQ1, a one-way analysis of covariance (ANCOVA) was conducted to compare the post-test argumentative writing performance of the experimental and control groups while controlling for pre-test scores (Field, 2024; Pallant, 2020). For RQ2, paired-samples t-tests were conducted to examine changes in writing self-efficacy from pre-test to post-test within each group, whereas independent-samples t-tests were conducted on gain scores to determine whether the improvement in writing self-efficacy differed significantly between the experimental and control groups.

The interview data were analysed using thematic analysis (Braun & Clarke, 2006). The Chinese transcripts were read repeatedly to develop familiarity with participants' accounts. Initial codes were generated inductively from the data to capture students' experiences of chatbot-supported writing instruction. Codes with related meanings were grouped into themes. A second researcher independently coded the transcripts, and coding discrepancies were resolved through discussion until consensus was reached. The qualitative findings were then integrated with the quantitative results to help explain the observed patterns in writing performance and self-efficacy.

Ethical considerations

All participants were informed of the objectives and procedures of the study before data collection, and participation was voluntary. Written informed consent was obtained from all participants. To protect confidentiality, participants' identities were replaced with identification codes, and all data were securely stored in password-protected files accessible only to the research team. Anonymity was maintained throughout the analysis and reporting of the findings.

RESULTS

Effects on argumentative writing performance

To address Research Question 1, descriptive statistics were first calculated to examine students' argumentative writing performance before and after the intervention. As shown in Table 1, both the experimental group (EG) and the control group (CG) improved from the pre-test to the post-test across all writing dimensions. However, the experimental group demonstrated consistently greater gains than the control group. In particular, the mean score for overall writing performance increased from 7.08 (SD = 0.94) to 7.82 (SD = 0.86) in the experimental group, whereas the control group improved from 7.14 (SD = 1.01) to 7.56 (SD = 1.01). Similar patterns were observed for content, structure, and language, suggesting that chatbot-supported process-genre writing instruction contributed to greater improvements in argumentative writing than conventional instruction.

Table 1 Descriptive statistics for pre-test and post-test argumentative writing performance

Writing Dimension	Control Group (n = 35)		Experimental Group (n = 34)	
	Pre-test (M)	Post-test (SD)	Pre-test (M)	Post-test (SD)
Content	7.30	1.11	7.80	1.14
Structure	6.99	0.97	7.34	0.91
Language	7.13	0.97	7.54	1.04
Overall writing performance	7.14	1.01	7.56	1.01

Before conducting the main analysis, the assumptions of ANCOVA were examined and found to be satisfactory. ANCOVA was then performed to compare the post-test scores of the two groups while controlling for pre-test performance. The results, presented in Table 2, revealed statistically significant group differences across all writing dimensions. Significant effects were found for content, structure, and language, indicating that the experimental group outperformed the control group after adjusting for baseline differences. A significant group effect was also observed for overall argumentative writing performance ($F = 50.09$, $p < .001$, partial $\eta^2 = .431$), representing a large effect size. Among the individual dimensions, the largest improvement was observed in language (partial $\eta^2 = .227$), followed by content (partial $\eta^2 = .199$) and structure (partial $\eta^2 = .139$). These

findings indicate that integrating a generative AI chatbot into the process-genre writing approach significantly enhanced students' argumentative writing performance, thereby providing support for Research Question 1.

Table 2 Results of ANCOVA for the experimental and control groups

Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Content	1.47	1	1.47	16.39	.000	.199
Structure	1.17	1	1.17	10.66	.000	.139
Language	2.13	1	2.13	19.37	.000	.227
Overall Writing Performance	1.72	1	1.72	50.09	.000	.431

Improvement in Writing Self-Efficacy

To examine changes in writing self-efficacy, paired-samples t-tests were first conducted separately for the experimental and control groups to compare pre-test and post-test scores on the overall scale and the five dimensions of writing self-efficacy: ideation, organization, grammar, English usage, and self-regulation. The results are presented in Table 3.

Table 3 Paired-samples t-test within the groups

Group	Measure	N	Pre-test Mean	Post-test Mean	MD	SD	t	df	p
EG	Total	34	3.12	4.54	1.42	0.81	10.15	33	<0.001
	Ideation	34	3.18	4.76	1.57	0.80	11.52	33	<0.001
	Organization	34	3.03	4.45	1.42	0.91	9.11	33	<0.001
	Grammar	34	3.04	4.32	1.28	1.04	7.14	33	<0.001
	English	34	2.95	4.42	1.47	0.94	9.15	33	<0.001
	Regulation	34	3.37	4.74	1.37	0.90	8.82	33	<0.001
CG	Total	35	3.21	3.90	0.69	0.09	45.06	34	<0.001
	Ideation	35	3.08	4.03	0.95	0.21	26.98	34	<0.001
	Organization	35	3.22	3.78	0.56	0.16	21.02	34	<0.001
	Grammar	35	3.18	3.89	0.71	0.23	18.13	34	<0.001
	English	35	3.19	3.77	0.58	0.21	16.45	34	<0.001
	Regulation	35	3.33	4.06	0.73	0.17	25.16	34	<0.001

For the experimental group, statistically significant improvements from pre-instructional questionnaire to post-instructional questionnaire were observed across all dimensions. Specifically, overall writing self-efficacy increased from 3.12 at pre-instructional questionnaire to 4.54 at post-instructional questionnaire (Mean difference = 1.42, $p < 0.001$). Similar highly significant improvements were found in all dimensions of self-efficacy including ideation, organization, grammar, English, and regulation.

In the control group, growth was also detected across all dimensions of writing self-efficacy. Overall writing efficacy increased from 3.21 to 3.90, with a mean difference of 0.69, $p < .001$. Furthermore, upward shifts were witnessed in ideation, organization, grammar, English, and regulation. While both groups exhibited enhancements over time, the experimental group generally demonstrated larger descriptive mean differences of 1.42 compared with a smaller increase in the control group with Mean difference at 0.69, suggesting that the AI-supported intervention was associated with stronger improvements in writing self-efficacy than conventional instruction.

To determine whether the experimental group demonstrated greater improvement in writing self-efficacy than the control group, independent-samples t-tests were subsequently conducted on the gain scores (post-test minus pre-test). The results are presented in Table 4.

Table 4 Independent t-test for experimental group and control group

Measure	EG (Mean Gain)	SD	CG (Mean Gain)	SD	t	df	p	Cohen's d
Total	1.42	0.81	0.69	0.09	5.25	67	<0.001	0.58
Ideation	1.57	0.80	0.95	0.21	4.48	67	<0.001	0.58
Organization	1.42	0.91	0.56	0.16	5.49	67	<0.001	0.65
Grammar	1.28	1.04	0.71	0.23	3.17	67	0.002	0.75
English	1.47	0.94	0.58	0.21	5.50	67	<0.001	0.67
Regulation	1.37	0.90	0.73	0.17	4.08	67	<0.001	0.65

The results showed that the experimental group demonstrated significantly greater improvement than the control group in overall writing self-efficacy (experimental group: M = 1.42, SD = 0.81; control group: M = 0.69, SD = 0.09, $p < .001$, d = 0.58). Significant between-group differences were also observed for ideation (experimental group: M = 1.57, SD = 0.80; control group: M = 0.95, SD = 0.21, $p < .001$, d = 0.58), organization (experimental group: M = 1.42, SD = 0.91; control group: M = 0.56, SD = 0.16, $p < .001$, d = 0.65), grammar (experimental group: M = 1.28, SD = 1.04; control group: M = 0.71, SD = 0.23, p = .002, d = 0.75), English usage (experimental group: M = 1.47, SD = 0.94; control group: M = 0.58, SD = 0.21, $p < .001$, d = 0.67), and self-regulation (experimental group: M = 1.37, SD = 0.90; control group: M = 0.73, SD = 0.17, $p < .001$, d = 0.65). These findings indicate that students who received the chatbot-supported process-genre writing approach demonstrated greater improvement in writing self-efficacy than those who received conventional instruction across the overall scale and all five dimensions.

Students' Experiences of Chatbot-Supported Process-Genre Writing Approach

To address Research Question 3, the interview data were analysed using thematic analysis. The analysis generated two overarching themes representing students' experiences of chatbot-supported process-genre writing instruction. The first theme, benefits gained, reflects the perceived pedagogical and psychological advantages of integrating the generative AI chatbot into the writing process. The second theme, challenges faced, captures the concerns and limitations students encountered while using the chatbot. Together, these themes comprised six codes, as summarised in Table 5.

Table 5. Themes and codes identified from the thematic analysis

Theme	Code	Summary of findings
Benefits gained	Writing process scaffolding	The chatbot supported planning, idea generation, organization, revision, language editing, and reflection throughout the writing process.
	Efficiency and accessibility	Students valued the chatbot's immediate responses, continuous availability, and rapid feedback beyond classroom hours.
	Affective transformation	The chatbot reduced writing anxiety, increased confidence, enhanced motivation, and encouraged greater engagement in writing tasks.
Challenges faced	Over-reliance	Students expressed concern that excessive dependence on the chatbot could reduce independent thinking and problem-solving.
	Proficiency mismatch barrier	AI-generated language was sometimes beyond students' English proficiency, making suggestions difficult to understand or apply.
	Homogenization and loss of individual voice	Students felt that excessive reliance on AI could reduce originality and make their writing increasingly similar to that of other learners.

Theme 1: Benefits gained

The first theme highlights the educational and affective benefits students associated with chatbot-supported process-genre writing instruction. Across the interviews, participants consistently described the chatbot as more than a tool for generating answers; instead, they viewed it as an instructional scaffold that supported their learning throughout the writing process while also making writing more efficient and psychologically less demanding. Three closely related codes emerged within this theme: writing process scaffolding, efficiency and accessibility, and affective transformation.

The most prominent code was writing process scaffolding. Students reported that the chatbot provided continuous support across different stages of argumentative writing, including planning, drafting, revising, editing, and reflection. During the planning stage, participants used the chatbot to generate ideas, organise essay structures, and develop outlines before beginning their drafts. As Student E explained, "*It is like asking the AI to provide a basic framework for the essay*". During revision, the chatbot helped students examine issues from multiple perspectives, strengthen counterarguments, and improve the logical organisation of their essays. Student B commented that "*AI provided both sides of the issue, helping me develop more balanced and well-supported arguments*", while Student C noted that "*AI helped reorganize the essay to improve its logic and coherence"*. Participants further explained that the chatbot assisted with language refinement by improving grammar, vocabulary, and expression during the editing stage. Student D stated that "*the tool was effective in improving my Chinese-style expressions"*, whereas Student F observed that "*it helped check word choice and grammar mistakes*". Taking all these findings into account, it can be understood that students in this study perceived the chatbot as providing integrated instructional support throughout the writing process, rather than functioning merely as a text-generation tool.

The second code concerned efficiency and accessibility. Students in the interviews repeatedly emphasised that the chatbot made the writing process more efficient because it provided immediate responses whenever support was needed. Unlike teacher or peer feedback, which rely on classroom time and availability, the chatbot offered continuous assistance that enabled students to revise their essays at their own pace. For example, Student F pointed out that "*AI allows me to receive immediate feedback at any time, even late at night*", while Student C highlighted its speed by stating that "*AI can respond very fast"*. Thus, the interviews showed that participants in the experimental group believed that this constant accessibility encouraged them to spend more time revising and refining their work, thereby extending learning beyond the classroom.

The third code, affective transformation, reflects the positive emotional changes students experienced while using the chatbot. Many participants explained that interacting with AI reduced the anxiety often associated with argumentative writing because they could ask questions repeatedly without fear of being judged. For example, Student E remarked that "*AI made me feel less anxious because I could ask simple questions without fear of judgment or impatience*", whereas Student G explained, "*No matter how weak my writing is, AI still responds patiently and tries to help me. This really motivates me to become a better writer*". Students also associated improvements in the quality of their writing with increased confidence in their own abilities. As Student H stated, "*AI corrected my grammar and word choice many times, so now I feel more confident when I read my own essays*". These findings indicate that students perceived the chatbot as supporting not only the cognitive demands of argumentative writing but also the affective factors that contribute to successful writing development.

Theme 2: Challenges faced

Despite recognising the educational value of chatbot-supported writing instruction, participants also identified some challenges that limited its effectiveness. Three interrelated codes emerged within this theme: over-reliance, the proficiency mismatch barrier, and homogenization and loss of individual voice. These concerns suggest that although students valued AI support in writing development, they also recognised the importance of maintaining independent thinking, matching AI assistance to learners' proficiency levels, and preserving their own authorial identity.

The most frequently reported concern was over-reliance. Participants acknowledged that the convenience and speed of the chatbot sometimes encouraged them to depend on AI rather than solving writing problems

independently. Some students worried that habitual reliance on AI could gradually weaken their critical thinking and reduce opportunities to develop their own writing strategies. Student I admitted, "*I started to rely on it too much*", while Student C observed that "*AI can make you stop thinking. If you rely on it for everything, you just wait for answers instead of thinking by yourself*". These comments indicate that students regarded AI as valuable support but recognised that excessive dependence on it could undermine the learning process.

The second code concerned the proficiency mismatch barrier. Participants explained that the chatbot frequently produced vocabulary, sentence structures, and expressions that exceeded their English proficiency, making some suggestions difficult to understand or incorporate into their writing. Instead of reducing cognitive demands, these responses occasionally increased confusion and frustration. Student B explained that "*AI-generated sentences were often too advanced"*, while Student C commented, "*Sometimes AI gives answers that are beyond my level. I just want something suitable for CET-4, but the vocabulary and sentence patterns it provides are more like CET-6*". These findings suggest that AI support is most effective when it is appropriately aligned with learners' language proficiency.

The final code was homogenization and loss of individual voice. Although students acknowledged that AI improved the overall quality and accuracy of their essays, many were concerned that repeated reliance on chatbot-generated suggestions could make their writing increasingly standardised. Participants felt that AI often recommended predictable sentence structures and conventional rhetorical patterns, making it more difficult to express personal ideas and writing style. For example, Student D explained that "*Many of AI's suggestions sound too standard and safe. They are correct, but they do not really show my own ideas or personal style"*, while Student F added, "*If I rely on AI too much, my writing may become less personal and sound like everyone else's"*. These findings highlight a central tension in AI-assisted writing: while chatbot support can improve writing quality, its educational value ultimately depends on learners' ability to critically evaluate AI-generated suggestions and preserve their own authorial voice.

Overall, the qualitative findings complement the quantitative results by demonstrating that students perceived chatbot-supported process-genre writing instruction as enhancing both the cognitive and affective dimensions of argumentative writing. At the same time, the findings highlight the importance of teacher mediation to encourage critical AI use, prevent over-reliance, provide guidance that matches learners' proficiency levels, and support the development of independent and authentic academic writers.

DISCUSSION

The findings indicate that embedding a generative AI chatbot within the process-genre approach significantly improved Chinese EFL students' argumentative writing performance and writing self-efficacy. The improvements in writing quality are consistent with previous studies showing that generative AI enhances idea development, organization, revision, language use, and overall writing performance (e.g., Can, 2026; Dong, 2024; Gao et al., 2025; Lai et al., 2026; Li & Tekwa, 2025; Luo et al., 2025; Pratama et al., 2025). Likewise, the gains in writing self-efficacy support recent intervention studies reporting that AI-assisted writing instruction can strengthen learners' confidence, motivation, and engagement (e.g., Barrot, 2023; Gao et al., 2025; Lai et al., 2026; Li & Tekwa, 2025; Pellas, 2023; Su et al., 2023; Zhang et al., 2025; Zhou & Wang, 2026). Unlike previous studies that examined AI primarily as a writing assistant or feedback provider, the present study embedded the chatbot within a structured process-genre framework and conceptualised it as a pedagogical scaffold that supported learners throughout the writing process. Viewed through the cognitive process theory of writing (Flower & Hayes, 1981) and social cognitive theory (Bandura, 2012), this instructional design reduced cognitive demands while promoting mastery experiences, constructive feedback, and supportive learning interactions. It may also explain why the present findings differ from studies reporting declines in writing self-efficacy following AI-supported instruction (Zhang et al., 2025).

The qualitative findings further explain these results by showing that students viewed the chatbot as a process-oriented scaffold that supported planning, drafting, revision, and reflection through immediate and accessible feedback. These findings reinforce recent research highlighting the collaborative and interactive role of generative AI in writing (e.g., Dong, 2024; Oppenheimer et al., 2025) while extending this body of research by

illustrating how AI can support both the cognitive and affective dimensions of writing when embedded within a structured pedagogical framework. Nevertheless, students also reported concerns about over-reliance, proficiency mismatch, and the potential loss of individual voice, echoing previous findings on learner dependency, proficiency differences, responsible AI use, and writing homogenization (e.g., Almusharraf et al., 2025; Nugroho et al., 2024; Yan, 2023). Collectively, these findings suggest that the educational value of generative AI depends less on the technology itself than on its pedagogically guided integration within established instructional approaches.

Despite its contributions, this study has some limitations. The participants were drawn from a single Chinese university, the intervention lasted only eight weeks, and the focus was limited to argumentative writing, which may restrict the generalizability of the findings. Future research should include larger and more diverse samples, adopt longitudinal and multi-institutional designs, and examine the use of generative AI across different writing genres and proficiency levels. Further studies could also investigate how varying levels of teacher mediation and AI support influence writing performance, writing self-efficacy, and the development of independent writing skills.

Practical Implications

The findings of this study have some important practical implications for EFL writing instruction. First, rather than using generative AI chatbots merely as writing assistants or feedback tools, instructors should integrate them systematically within established pedagogical approaches such as the process-genre approach, where AI provides stage-specific support for planning, drafting, revising, editing, and reflection. Second, the significant improvements in writing self-efficacy suggest that AI should be used to complement teacher guidance and peer interaction, creating a supportive learning environment that promotes both cognitive and affective development. Third, the qualitative findings indicate that teacher mediation remains essential to help students critically evaluate AI-generated suggestions, prevent over-reliance, match AI support to learners' proficiency levels, and preserve students' authorial voice. These findings suggest that the educational value of generative AI depends not only on the technology itself but also on its pedagogically guided integration into writing instruction.

CONCLUSION AND LIMITATIONS

This study examined the effectiveness of a generative chatbot-supported process-genre writing approach in enhancing Chinese university EFL students' argumentative writing performance, writing self-efficacy, and learning experiences in a blended learning environment. The findings indicate that integrating generative AI within the process-genre approach improved students' writing performance and strengthened all dimensions of writing self-efficacy. Students also viewed the chatbot as a pedagogical scaffold that supported different stages of the writing process, although concerns regarding over-reliance, proficiency mismatch, and the preservation of individual voice remained. More importantly, the study extends current AI-assisted writing research by demonstrating that the educational value of generative AI lies in its systematic integration within an established pedagogical framework rather than its role as a standalone writing tool. By conceptualising the chatbot as a pedagogical scaffold, the study provides a clearer understanding of how generative AI can support both the cognitive and affective dimensions of EFL writing in blended learning contexts.

A key limitation of this study is that each condition involved only one intact class taught by a different instructor. Consequently, class characteristics and teaching practices may have contributed to the observed differences between groups. Although the instructors received training and pre-test scores were statistically controlled, these measures could not resolve the confounding of instructional condition with class and instructor, limiting causal interpretation of the findings. Furthermore, the comparison between chatbot-supported process-genre instruction and conventional instruction did not allow the chatbot’s contribution to be distinguished from other components of the program. Future research should include multiple classes and instructors per condition and compare process-genre instruction with and without chatbot support.

REFERENCES

1. Almusharraf, A., Bailey, D., Almusharraf, N., & Alotaibi, T. (2025). Students' perceptions of generative AI in EFL writing: Strategies, self-efficacy, satisfaction and behavioural intention. Australasian Journal of Educational Technology, 41(5), 18-36. https://doi.org/10.14742/ajet.10045
2. Awada, G., Burston, J., & Ghannage, R. (2020). Effect of student team achievement division through WebQuest on EFL students' argumentative writing skills and their instructors' perceptions. Computer Assisted Language Learning, 33(3), 275-300. https://doi.org/10.1080/09588221.2018.1558254
3. Baidu, Inc. (2023, March 16). Baidu unveils ERNIE Bot, the latest generative AI mastering Chinese language and multi-modal generation [Press release]. PR Newswire. https://www.prnewswire.com/news-releases/baidu-unveils-ernie-bot-the-latest-generative-ai-mastering-chinese-language-and-multi-modal-generation-301774240.html
4. Bandura, A. (1997a). Self-efficacy in changing societies. Cambridge university press.
5. Bandura, A. (1997b). Self-efficacy: The exercise of control. Worth Publishers.
6. Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. Journal of Management, 38(1), 9-44. https://doi.org/10.1177/0149206311410606
7. Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. Assessing Writing, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745
8. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa
9. Can, Y. (2026). Artificial intelligence vs. instructor supported AI: Teaching second language writing via "Write&Improve". Interactive Learning Environments, 34(3), 1154-1177. https://doi.org/10.1080/10494820.2025.2519123
10. Chang, J. L., Hung, H. T., & Yang, Y. T. C. (2024). Effects of an annotation-supported Socratic questioning approach on students' argumentative writing performance and critical thinking skills in flipped language classrooms. Journal of Computer Assisted Learning, 40(1), 37-48. https://doi.org/10.1111/jcal.12856
11. Chen, H.-L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. Journal of Educational Computing Research, 58(6), 1161-1189. https://doi.org/10.1177/0735633120929622
12. Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. Journal of Research in Science Teaching, 47(7), 883-908. https://doi.org/10.1002/tea.20385
13. Dong, L. (2024). "Brave New World" or not?: A mixed-methods study of the relationship between second language writing learners' perceptions of ChatGPT, behaviors of using ChatGPT, and writing proficiency. Current Psychology, 43(21), 19481-19495. https://doi.org/10.1007/s12144-024-05728-9
14. Ebadi, S., & Amini, A. (2024). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. Interactive Learning Environments, 32(2), 655-673. https://doi.org/10.1080/10494820.2022.2096638
15. Field, A. (2024). Discovering statistics using IBM SPSS statistics. Sage.
16. Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition & Communication, 32(4), 365-387. https://doi.org/10.58680/ccc198115885
17. Gao, J., Zhang, J., & Li, Y. (2025). Do AI Chatbot-integrated writing tasks influence writing self-efficacy and critical thinking ability? A quasi-experimental study. Computers and Education: Artificial Intelligence, 9, 1-10. https://doi.org/10.1016/j.caeai.2025.100472
18. Grigg, S., Perera, H. N., McIlveen, P., & Svetleff, Z. (2018). Relations among math self efficacy, interest, intentions, and achievement: A social cognitive perspective. Contemporary Educational Psychology, 53, 73-86. https://doi.org/10.1016/j.cedpsych.2018.01.007
19. Guo, K., Li, Y., Li, Y., & Chu, S. K. W. (2024). Understanding EFL students' chatbot-assisted argumentative writing: An activity theory perspective. Education and Information Technologies, 29(1), 1-20. https://doi.org/10.1007/s10639-023-12230-5
20. Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. Computers & Education, 203, 104862. https://doi.org/10.1016/j.compedu.2023.104862

21. Han, D.-E. (2020). The effects of voice-based AI chatbots on Korean EFL middle school students' speaking competence and affective domains. Asia-pacific Journal of Convergent Research Interchange, 6(7), 71-80. https://doi.org/10.47116/apjcri.2020.07.07
22. Hartwell, K., & Aull, L. (2022). Constructs of argumentative writing in assessment tools. Assessing Writing, 54. https://doi.org/10.1016/j.asw.2022.100675
23. Hayes, J. R. (2012). Modeling and remodeling writing. Written Communication, 29(3), 369-388. https://doi.org/10.1177/0741088312451260
24. Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. Journal of Computer Assisted Learning, 38(1), 237-257. 10.1111/jcal.12610
25. Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. Computer Assisted Language Learning, 36(7), 1338-1364. https://doi.org/10.1080/09588221.2021.1987272
26. Jin, S.-H. (2025). Measures of learner-generative AI relationships. Computers and Education Open, 8, 100258. https://doi.org/10.1016/j.caeo.2025.100258
27. Kerman, N. T., Noroozi, O., Banihashem, S. K., Karami, M., & Biemans, H. J. (2024). Online peer feedback patterns of success and failure in argumentative essay writing. Interactive Learning Environments, 32(2), 614-626. https://doi.org/10.1080/10494820.2022.2093914
28. Kohnke, L. (2023). A pedagogical chatbot: A supplemental language learning tool. RELC Journal, 54(3), 828-838. https://doi.org/10.1177/00336882211067054
29. Kong, A., & Teng, M. F. (2023). The operating mechanisms of self-efficacy and peer feedback: An exploration of L2 young writers. Applied Linguistics Review, 14(2), 297-328. https://doi.org/10.1515/applirev-2020-0019
30. Kuhn, D., & Moore, W. (2015). Argumentation as core curriculum. Learning: Research and Practice, 1(1), 66-78. https://doi.org/10.1080/23735082.2015.994254
31. Lai, C., Pan, M., Guo, K., & Cui, Y. (2026). What task to offload to GenAI in the writing feedback process?–Effects of task-offloading approaches on EFL learners' writing skill development. Computers & Education, 252, 105675.
32. Li, H., Wang, Y., Luo, S., & Huang, C. (2025). The influence of GenAI on the effectiveness of argumentative writing in higher education: Evidence from a quasi-experimental study in China. Journal of Asian Public Policy, 18(2), 405-430. https://doi.org/10.1080/17516234.2024.2363128
33. Li, M., & Tekwa, K. (2025). Timing AI intervention in the writing process for low proficiency learners: relationship with motivation, self-efficacy, and performance: Timing AI intervention in the writing process for low proficiency learners: relationship with motivation, self-efficacy, and performance. European Journal of Psychology of Education, 40(4), 1-35. https://doi.org/10.1007/s10212-025-01030-9
34. Li, Z., & Yan, D. (2020). Effect of pigai. org on english majors' writing self-efficacy and writing performance. Journal of Physics: Conference Series,
35. Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing. Education and Information Technologies, 30(12), 17695-17715. https://doi.org/10.1007/s10639-025-13488-7
36. Mahdi, H. S., Milad, M., & Saeed, A. T. (2025). Evaluating the effectiveness of AI tools across the essay writing process: a comparative study of pre-writing, drafting, and post-writing stages. International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 15(1), 1-20. https://doi.org/10.4018/IJCALLT.391354
37. Nugroho, A., Putro, N. H. P. S., Syamsi, K., Mutiaraningrum, I., & Wulandari, F. D. (2024). Teacher's experience using ChatGPT in language teaching: An exploratory study. Computers in the Schools, 1-20. https://doi.org/10.1080/07380569.2024.2441161
38. OpenAI. (2024). A student's guide to writing with ChatGPT. https://chatgpt.com/use-cases/student-writing-guide/
39. Oppenheimer, D. M., Cash, T. N., & Connell Pensky, A. E. (2025). You've got AI friend in me: LLMs as collaborative learning partners. International Journal of Artificial Intelligence in Education, 35(6), 3896-3921. https://doi.org/10.1007/s40593-025-00521-6

40. Pagano, N., Bernhardt, S. A., Reynolds, D., Williams, M. T., & McCurrie, K. (2008). An inter-institutional model for college writing assessment. College Composition & Communication, 60(2), 285-320. https://doi.org/10.58680/ccc20086868
41. Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. Reading &Writing Quarterly, 19(2), 139-158. https://doi.org/10.1080/10573560308222
42. Pallant, J. (2020). SPSS survival manual: A step by step guide to data analysis using IBM SPSS. Routledge.
43. Pellas, N. (2023). The effects of generative AI platforms on undergraduates' narrative intelligence and writing self-efficacy. Education Sciences, 13(11), 1155. https://doi.org/10.3390/educsci13111155
44. Pratama, R. D., Widiati, U., & Hakim, L. N. (2025). Effects of human-AI collaborative writing strategy on EFL students' argumentative writing skills. Computer Assisted Language Learning, 1-32. https://doi.org/10.1080/09588221.2025.2503900
45. Ruwe, T., & Mayweg-Paus, E. (2023). "Your argumentation is good", says the AI vs humans–The role of feedback providers and personalised language for feedback effectiveness. Computers and Education: Artificial Intelligence, 5, 100189. https://doi.org/10.1016/j.caeai.2023.100189
46. Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. Contemporary Educational Psychology, 60, 101832. https://doi.org/10.1016/j.cedpsych.2019.101832
47. Shen, X., & Tao, Y. (2025). Metacognitive strategies, AI-based writing self-efficacy and writing anxiety in AI-assisted writing contexts: A structural equation modeling analysis. International Journal of TESOL Studies, 7(1), 70-87. https://doi.org/10.58304/ijts.20250105
48. Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. Assessing Writing, 57, 1-11. https://doi.org/10.1016/j.asw.2023.100752
49. Su, Y., Liu, K., Lai, C., & Jin, T. (2021). The progression of collaborative argumentation among English learners: A qualitative study. System, 98, 1-15. https://doi.org/10.1016/j.system.2021.102471
50. Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. System*, 90*, Article 102221. https://doi.org/10.1016/j.system.2020.102221
51. Teng, L. S., Sun, P. P., & Xu, L. (2018). Conceptualizing writing self-efficacy in English as a foreign language contexts: Scale validation through structural equation modeling. Tesol Quarterly, 52(4), 911-942. https://doi.org/10.1002/tesq.432
52. Wang, Y., Wang, Y., Pan, Z., & Ortega-Martín, J. L. (2024). The predicting role of EFL students' achievement emotions and technological self-efficacy in their technology acceptance. The Asia-Pacific Education Researcher, 33(4), 771-782. https://doi.org/10.1007/s40299-023-00750-0
53. Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. Frontiers in Psychology, 14, 1249991. https://doi.org/10.3389/fpsyg.2023.1249991
54. Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. System, 39(4), 510-522. https://doi.org/10.1016/j.system.2011.10.017
55. Xue, J., & Chen, M. (2026). Unravelling the relationship between self-efficacy and self-regulated strategies in GenAI-assisted writing of postgraduates: the mediating role of feedback engagement. Frontiers in Psychology, 17, 1-12. https://doi.org/10.3389/fpsyg.2026.1786221
56. Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. Education and Information Technologies, 28(11), 13943-13967. https://doi.org/10.1007/s10639-023-11742-4
57. Yang, Y. (2020). Engagement Resources in Chinese College Students' Argumentative Writings. In B. Yang & W. Li (Eds.), Corpus-based approaches to grammar, media and health Discourses: Systemic functional and other perspectives (pp. 251-281). Springer.
58. Yao, G., & Fan, L. (2026). Self-efficacy in AI-assisted L2 academic writing: scale development and validation. BMC Psychology, 14(1), 1-16. https://doi.org/10.1186/s40359-026-04503-8
59. Zhang, R., Zou, D., & Cheng, G. (2023). Chatbot-based training on logical fallacy in EFL argumentative writing. Innovation in Language Learning and Teaching, 17(5), 932-945. https://doi.org/10.1080/17501229.2023.2197417

60. Zhang, R., Zou, D., Cheng, G., & Xie, H. (2025). Flow in ChatGPT-based logic learning and its influences on logic and self-efficacy in English argumentative writing. Computers in Human Behavior, 162, 108457. https://doi.org/10.1016/j.chb.2024.108457
61. Zhang, S., Shan, C., Lee, J. S. Y., Che, S., & Kim, J. H. (2023). Effect of chatbot-assisted language learning: A meta-analysis. Education and Information Technologies, 28(11), 15223-15243.
62. Zhou, C., & Wang, Y. (2026). University students' writing feedback literacy in the AI era: The interplay of generative AI acceptance, writing anxiety and writing self-efficacy in Chinese EMI educational contexts. European Journal of Education, 61(1), e70478. https://doi.org/10.1111/ejed.70478
63. Zhu, W. (2001). Performing argumentative writing in English: Difficulties, processes, and strategies. TESL Canada Journal, 34-50. https://doi.org/10.18806/tesl.v19i1.918